AD_____

Award Number: DAMD17-99-1-9273

TITLE: The Regulatory Interactions of p21 and PCNA in Human Breast Cancer

PRINCIPAL INVESTIGATOR: Derek J. Hoelz
Linda H. Malkas, Ph.D.

CONTRACTING ORGANIZATION: University of Maryland
Baltimore, Maryland 21201

REPORT DATE: July 2001

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2001	Annual Summary (1 Jul 00 - 30 Jun 01)

4. TITLE AND SUBTITLE
The Regulatory Interactions of p21 and PCNA in Human Breast Cancer

5. FUNDING NUMBERS
DAMD17-99-1-9273

6. AUTHOR(S)
Derek J. Hoelz
Linda H. Malkas, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Maryland
Baltimore, Maryland 21201

E-Mail: dhoelz@som.umaryland.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. Abstract *(Maximum 200 Words)* *(abstract should contain no proprietary or confidential information)*

Proliferating cell nuclear antigen (PCNA) is a multifunctional protein essential for DNA replication and DNA repair. Recently, our laboratory has shown that when resolved using two-dimensional polyacrylamide gel electrophoresis (2D PAGE) PCNA focuses at distinct isoelectric point(s) (pI). It was subsequently noted that this focusing pattern was different when non-malignant and malignant breast cells were compared. The PCNA present in non-malignant breast cells focuses at one pI on 2D PAGE while the PCNA present in malignant breast cells consistently focuses at two distinct pIs. In order to explore the functional consequences of the alteration of PCNA in malignant breast cells, we have begun to examine its interaction with cell cycle inhibitor $p21^{WAF1/CIP1/SDI1}$. Through its interaction with PCNA, p21 inhibits DNA replication in response to DNA damage. Theoretically, p21 stops the cell from replicating damaged DNA and allocates time needed for DNA repair. Paradoxically, PCNA also functions in repair, and there is a multitude of conflicting data on the inhibitory role of p21 in DNA repair. Therefore, we have begun to elucidate the structure and location of the post-translational modification on PCNA and have shown that p21 differentially interacts with PCNA.

14. Subject Terms (keywords previously assigned to proposal abstract or terms which apply to this award)
Breast Cancer MCF7, p21, PCNA, synthesome, DNA replication, DNA fidelity, cell

15. NUMBER OF PAGES
81

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

To better understand the role of DNA replication in breast cancer, it is essential to examine the machinery that carries out the DNA synthetic process. Our laboratory has successfully purified a complex of proteins from breast cells that is fully competent to carry out T-antigen dependent, SV40 origin specific DNA replication *in vitro*, which we have termed the DNA synthesome (Malkas *et al.*, 1990). Analysis of the constituent proteins of the DNA synthesome of malignant and non-malignant breast cells by two-dimensional polyacrylamide gel electrophoresis (2D PAGE) has uncovered a modification in an essential DNA replication protein, proliferating cell nuclear antigen (PCNA) (Bechtel *et al.*, 1998). 2D PAGE analysis revealed that PCNA of malignant breast cells resolves as two distinct species, a unmodified form and a modified form, while PCNA present in non-malignant cells resolves exclusively as one form, the modified form. PCNA functions by forming a trimeric "sliding clamp" that encircles the DNA and interacts with the DNA polymerases δ and ε (Krishna *et al.*, 1994; Zhang *et al.*, 1998; Eki, T.). Polymerase δ carries out leading strand DNA synthesis, and although a role for polymerase ε has not yet been ascribed, it has been hypothesized to function in DNA repair. Another protein that interacts with PCNA is $p21^{WAF1/CIP1/SDI1}$. P21 is a CDK inhibitor that, when induced by p53 in response to DNA damage binds PCNA and effectively competes away polymerase δ leading to the efficient inhibition of DNA replication. This inhibition impedes the replication of damaged DNA and theoretically allots time for the cell to repair its damaged DNA. Therefore, any alterations of the PCNA molecule could potentially abrogate p21binding leading to replication of damaged DNA and/or insufficient time for DNA repair. It is our goal to study the interaction of p21 with the modified and non-modified forms of PCNA and to investigate any functional consequences alterations in PCNA/p21 binding may have on DNA replication, DNA repair, and DNA replication fidelity.

Progress Report

Purification and separation of the two forms of PCNA

Two dimensional polyacrylamide gel electrophoresis (2D PAGE) and electroelution of the different forms of PCNA in MCF7 cells was performed in an attempt to purify the two forms of PCNA. Initially, MCF7 cells were fractionated to a clarified nuclear extract and post-microsomal (NE/S3) fraction and placed onto a Phosphocellulose column. Phosphocellulose dissociated the synthesome multiprotein complex and resolves PCNA from the DNA polymerases and other synthesome associated DNA replication proteins. The column eluate fraction containing PCNA was then passed over a hydrophobic Phenyl Sepharose column followed by anion exchange chromatography (Q Sepharose). Purified PCNA was then loaded onto 2D PAGE and resolved. One of the gels was silver stained and analyzed using the Melanie II software. Analysis predicted the general location of the PCNA species and empirically predicted the isoelectric points (pIs) of the two species. The pIs for PCNA were 4.74 for the unmodified (cancer specific) form, which is close to the theoretical pI for PCNA (4.56), and 6.96 for the modified (non-malignant) form. The unmodified form and modified form of PCNA were then excised from four parallel 2D PAGE gels and electroeluted into 25mM HEPES buffer pH 7.2 using a Mini Whole-Gel Eluter (Bio-Rad). The presence of PCNA in the electroeluted fractions was determined by slot blotting (Life Technologies, Inc.), and the fractions containing PCNA were subsequently tested by 2D PAGE. The 2D gels confirmed the presence of the unmodified form in the electroelute; however, 2D analysis of the modified form of PCNA in the electroelute demonstrated the exclusive presence of the cancer-specific form. It was concluded from these data that the modification of PCNA was lost in either the electroelution or, more probably, the second dimension SDS PAGE step. Because of the apparent loss of modification, new approaches to separating the two forms of PCNA have begun to be developed.

In order to separate the two forms of PCNA, an alternative chromatographic protocol was employed. The Q-Sepharose chromatography has been replaced with cation exchange chromatography (SP Sepharose). PCNA eluted off the Phenyl Sepharose column will be dialyzed into 50mM potassium phosphate buffer with a pH of 3.0. Because the pH of the buffer is below the pIs for the two forms of PCNA, the molecules will have a positive character and bind to the SP Sepharose. The unmodified form can now be eluted from the column in 50mM potassium phosphate pH 5.75, and because the pH of this buffer is still below the pI of the modified form, it will still bind the column and not be eluted. The modified form can then be eluted into 50mM potassium phosphate pH 8.0. Although this protocol was somewhat effective at separating the two forms of PCNA, an easier and more consistent method was subsequently developed utilizing a salt gradient rather than a pH shift to eluted PCNA from the SP Sepharose column. In addition, the pH of the buffer was changed to 50mM malonic acid containing 1mM DTT, pH 5.8. This new protocol allowed for the cancer-specific form to flow through the column, while binding the non-malignant form. Elution of the column into a linear NaCl gradient up to 0.5M yielded a essentially homogenous solution of the nonmalignant form of PCNA. The flow through fraction containing the cancer-specific form of PCNA was then further purified by passage over a Q-Sepharose column in 50mM Tris-HCl, 0.1M KCl, 1mM DTT, pH 9.0, followed by elution into a linear KCL gradient up to 1M. Resolution of the cancer specific form of PCNA by 2D-PAGE isolated using this protocol demonstrated a single form of PCNA that focuses at 4.7 pI, however,

problems have been encountered while trying to confirm the pI of the non-malignant form of PCNA isolated off of the SP Sepharose column. Initial identification of fractions containing PCNA were accomplished by slot blotting/ Western blotting with anti-PCNA (PC10, Oncogene Sciences, Inc.)antibodies, but the PCNA concentrations of the fractions were too dilute and contained too high of a salt concentration to be effectively resolved on an isoelectric gel. Manipulation of the samples in effort to increase the protein concentration was attempted numerous times using numerous available methods with little effectiveness. Due to the strong hydrphobicity of the protein, ultrafiltration yielded a lower protein concentration than that of the staring material. Speed-vacuum lyophilization was also attempted, and the protein fell out of solution and was not recoverable even with the highly denaturing conditions of the 2D-PAGE sample buffer. Precipitation with NH_3SO_4 had similar results.The major issue was determined to be solubility. Dialysis and freeze-thawing of the sample was shown to precipitate the majority of the purified non-malignant form of PCNA, which most probably due to PCNA's strong hydrophobicity and it's propensity to from multimers. Interestingly, isolation of the cancer-specific form does not seem to readily form aggregates when purified. Use a different buffer, 20mM citrate, pH 4.8 and 20% v/v ethylene glycol during the SP Sepharose chromatography in combination with dilution and size exclusion chromatography using G-25 beads in attempt to change the buffers pH while stabilizing the non-malignant form of PCNA in solution have been performed with increasing success. However, confirmation of the proper pI for the purified non-malignant form of PCNA must still be completed.

Cloning PCNA into a protein expression vector

In order to study PCNA and its modification(s) we are going to express and modify the protein *in vitro*. To do this we cloned the PCNA from MCF7 cells. The cDNA sequence encoding the entire PCNA translation unit was prepared using Superscript Preamplification System First Strand DNA Synthesis (Life Technologies, Inc.) followed by cloning of the amplified cDNA into the vector pCR2.1 (InVitrogen) according to the manufacturers instructions. Total RNA was isolated using Trizol reagent (Life Technologies, Inc.). Second strand DNA synthesis was carried out by priming the first strand cDNA with oligonucleotide 5'-GCGTTGTTGCCACTCCGC-3' on the 5' end of the cDNA and 5'-GCAGTTCTCAAAGAGCTTAG-3' on the 3' end of the cDNA and amplifying the primed first strand using reverse transcriptase PCR.

Subcloning of the cDNA was done using the pDUAL expression and cloning kit according to the manufacturer's protocol (Stratagene). Briefly, Eam1104I restriction sites were engineered onto the 5' ends of the PCR primers 5'-ATGTTCGAGGCGCGCCTGGTCCAG-3' and 5'-AGATCCTTCATCCTCGATCTTGGGAGC-3' and the amplified PCNA cDNA was inserted into the pDUAL expression vector. Purification of PCNA is accomplished using Calmodulin Affinity Resin (Stratagene) which specifically binds a Calmodulin Binding Protein (CBP) tag fused to the C-terminus of the protein.

2D PAGE, tryptic digestion of PCNA, MALDI-TOF, and ESI-MS/MS mass spectrometry

A new approach has been employed to identify the amino acid sequence harboring the PCNA modification. MCF7 cells were fractionated and PCNA was purified through a Q Sepharose chromatography as outlined in above. The two forms of PCNA were then resolved on 2D PAGE

and visualized by silver staining using Silver Stain Plus (Bio-Rad). The spots corresponding to PCNA are then excised from the gel and digested with trypsin. The resultant peptide fragments are then resolved using an MALDI-TOF mass spectrometer in the negative ion, reflectron mode. The molecular masses of the peptide fragments obtained by MALDI-TOF mass spectrometry are then used to search protein sequence databases available on the inter-net. Sites of post-translational modification can then be determined by increases in the apparent molecular masses of the peptide fragment(s). The mass difference of these shifted peptide fragments will give insight into the identity and location of the modification(s) on PCNA.

Using a Voyager DE STR MALDI-TOF mass spectrometer (PerSeptive Biosystems, Inc.) we were able to resolve peptide fragments derived form PCNA excised from 2D PAGE, and by searching the proteomics database Protein Prospector (UCSF), we were able to identify the protein as PCNA by its tryptic digest pattern. Further work will enable us to resolve the modified fragment(s) identifying the location and nature of the modification(s).

An additional approach to identification of PCNA structure and post-translational modification has been employed with promising results. Phosphocellulose chromatography and phenyl Sepharose chromatography followed by resolution on 2D-PAGE first isolated the two forms of PCNA. The 2D-PAGE gels were then stained with colloidal Coomassie (Pierce) and spots at or near the pI and molecular weight for the two forms of PCNA were removed and subjected to *in situ* digestion with sequencing grade trypsin (Promega). The extracted peptide fragments were then subjected to separation by reverse-phase HPLC and fed directly into a LCQ ion trap mass spectrometer (Finnagen) by electrospray ionization (ESI). Dissociation of the parent peptide by collision with an inert gas yielded peptide sequence data that allowed for protein identification using two different search algorithms, Mascot and Seaquest, which identified at least two distinct spots on the 2D-PAGE gels that correspond to PCNA. Analysis of the peptides (11 peptides were identified out of 21 that contain no missed tryptic cleavage sites) did not yield any post-translational modifications except for oxidized methionines. To better identify the possible protein modifications on PCNA, spots corresponding to PCNA will be digested *in situ* by two or more different proteolytic enzymes, trypsin, Lys C, or cyanogen bromide, which will allow for identification and analysis of a greater number of proteolytic fragments spanning a larger area of the PCNA protein ultimately aiding in identification of peptides/amino acids harboring post-translational modifications and their structure.

In vitro transcription and translation of GST-p21

To study the interactions of p21 with the two forms of PCNA, GST- p21 protein was produced *in vitro* using *E. coli* T7 S30 Extract System for Circular DNA (Promega) for use in Far Western blotting experiments (see below). GST-p21 was labeled by addition of ^{3}H-leucine to the *in vitro* transcription and translation reaction and purified on a Glutathione Sepharose column (Amersham Pharmacia). Presence of *in vitro* protein product was confirmed by SDS-PAGE and autoradiography.

Co-immune precipitation, GST pull-down assays, and Far Western blotting

To demonstrate that the modification(s) on PCNA effects its ability to interact with p21, three different experimental approaches were taken. First, the interaction of p21 with the two forms of PCNA was examined by co-immune precipitation. MCF7 cells were fractionated to a NE/S3 and incubated with monoclonal anti-p21 antibody (DF10, Oncogene Research) at 4°C for 2 hours. The antibodies were then bound to Protein-A-Agarose (Oncogene Research) and washed. The precipitate was then resolved by 2D PAGE, transferred to nitrocellulose, and Western blotted using anti-PCNA antibody (PC10, Oncogene Research) (Figure 4). Secondly, GST pull-down experiments were performed. GST-p21 was purified from inclusion bodies *E. coli* BL21 (DE-3) cells (Stratagene) and purified using Glutathione Sepharose. MCF7 NE/S3 was then added to the GST-p21 conjugated Glutathione Sepharose beads and incubated at 4°C for 2h. The beads were washed and loaded onto 2D PAGE. The 2D gels were transferred to nitrocellulose and Western blotted with PCNA antibody. The co-immune precipitation and GST pull-down experiments showed similar results. They both demonstrated the presence of PCNA; however, the precipitated PCNA has a pI in-between that of the two forms of PCNA found in the MCF7 NE/S3 control. Another assay used to elucidate binding of p21 to PCNA was Far Western blotting. Far Western blotting utilizes a Western blotting approach that uses a labeled p21 instead of PCNA antibody to detect PCNA. MCF7 NE/S3 was first resolved on 2D PAGE and transferred to PVDF membrane. The proteins immobilized on the membrane were then denatured in buffer containing 6M guanidine HCl for 1h, slowly refolded in 3M guanidine HCl and step-wise down to 0.187 M and finally into buffer. The re-natured membranes were incubated with labeled GST –p21 (see above) over night at 4°C. The membranes were washed three times with buffer, dried, sprayed with En^{3}Hance (New England Nuclear) and exposed to autoradiography at –80°C. Despite numerous attempts, we were unable to detect p21 binding to PCNA on 2D gels by Far Western blotting. Although spots are visible on the autoradiographs, indicating p21 binding, none of the spots overlap PCNA visualized by Western blotting of 2D gels run in parallel. Due to the inability of these experiments to demonstrate that the modification(s) on PCNA effects p21 binding, we have begun to take a new experimental direction.

BIACORE analysis of the interactions between p21 and the two forms of PCNA

To definitively show that the modification(s) on PCNA affects its ability to interact with p21, we have now begun to use BIAcore 2000 surface plasmon resonance (SPR) mass spectrometer (BIAcore, Inc). The cancer-specific and non-malignant forms of PCNA isolated as described above (Purification and separation of the two forms of PCNA) were bound to BIACORE CM-5 chips by carbidamide linkage. Next, varying concentrations of p21 were passaged over the chip surface. Binding events were detected by changes in molecular mass. Concentrations of p21 at 1000nM, 800nM, 400nM, 200nM, 100nM, and 50nM were passaged over the cancer-specific form of PCNA bound to the sensor chip at a flow rate of 100µl/min for 30s, and the binding reactions were subjected to analysis using BIAevaluation software version 2.1 (BIAcore, Inc.). The binding of p21 non-malignant form of PCNA was also analysed using the same reaction conditions although the concentrations of p21 were 678nM, 339nM, 113nM, 85.2nM, and 28.4nM. The K_d values for p21 binding events were 1.04×10^{-5} for the cancer-specific form of PCNA, and 1.07×10^{-10} for the non-malignant form of PCNA demonstrating a 10,000-fold greater affinity of p21 for the non-malignant form of PCNA as compared to the cancer-specific form of PCNA. Although the 1:1 Langmuir binding models assigned to the empirically derived

data fit the curves accurately enough to assume that any departure of the derived curves from the model is noise, further experimental data must be generated to definitively assign a K_d for the two interactions. According to Myszka *et al.* BIACORE kinetic data must performed in such a manner as to control for the effects of avidity and mass action. Therefore, the BIACORE experiments were performed by linking PCNA (and not p21) to the sensor chip in order to avoid any possible avidity effects PCNA may have on the surface of the chip through the forming multimers. Additionally, the flow rates were increased to 100 µl/min and the surface exposure to p21 was limited to 30s to limit mass action. However, the response units (RU) of the kinetic experiments should also be kept below 50 RUs in order to minimize mass action, and these experiments were preformed with 182 RUs for the cancer-specific form and 61 RUs for the non-malignant form and would therefore need to be repeated. Along with lowering the RUs for the reactions, competition experiments will also be performed in the presence of synthetic peptides derived from the p21 interacting interdomain connector loop as well as a non-specific region of PCNA that does not bind p21 in varying concentrations to demonstrate the specificity of the PCNA p21 interaction.

Key Research Accomplishments

- Performed 2D PAGE and electroeluted PCNA
- Tryptic digested PCNA excised from 2D PAGE and analyzed peptide fragments by MALDI-TOF mass spectrometry
- Cloned PCNA into pDUAL plasmid expression vector
- Co-immune experiments using anti-p21 antibody and analysis of precipitated PCNA by 2D PAGE
- Precipitation of PCNA by GST-p21 and analysis by 2D PAGE
- *In vitro* transcription, translation and purification of GST-p21
- Far Western blotting of 2D PAGE with ^{3}H labeled GST-p21
- Purified the cancer-specific and non-malignant forms of PCNA from MCF7 cells
- Identified PCNA sequence by ion-trap mass spectrometry
- Empirically derived Kd values for the cancer-specific and non-malignant forms of PCNA binding to p21 using BIACORE

Reportable Outcomes

Abstracts

Hoelz, D.J., Park, M., Dogruel, D., Bechtel, P., Sekowski, J., Xiang, H.Y., Hickey, R.J., Malkas, L.H. (2000): Analysis of a Malignant Cell's DNA Replication Apparatus by Mass Spectrometry. Scientific Proceeding of the 91st Annual Meeting of the American Association for Cancer Research. 41: 847.

Hoelz, D.J., Bechtel, P., Freund, R., Park, M., Hickey, R.J., Malkas, L.H. (2001) Isolation and Characterization of the Non-malignant Form of PCNA from MCF7 Breast Cancer Cells. Scientific Proceeding of the 92nd Annual Meeting of the American Association for Cancer Research. 42: 894.

Hoelz, D.J., Bechtel, P., Hickey, R.J., Malkas, L.H. (2001) Differential binding of p21WAF1 by an altered form of proliferating cell nuclear antigen present in breast cancer cells. 24th Annual San Antonio Breast Cancer Symposium.

Papers

Hoelz, D.J., Bechtel, P., Hickey, R.J., Malkas, L.H. (2001): Purification of the Malignant Form of Proliferating Cell Nuclear Antigen from Breast Cancer cells. Manuscript in preparation.

D. Hoelz, R. J. Hickey, and **L. H. Malkas,** Prokaryotic DNA Replication. (2001) *Encyclopedia of Life Sciences,* Nature Publishing (in press).

R. J. Hickey, D. Hoelz and **L. H. Malkas,** Eukaryotic DNA Replication. (2001) *Encyclopedia of Life Sciences,* Nature Publishing (in press).

Patents

Hoelz, D.J., Malkas, L.H., Hickey, R.J. (2001) The Isolation of Modified and Non-modified Forms of Proliferating Cell Nuclear Antigen (PCNA) for Diagnostic/ Drug development.

Conclusions

- **The unmodified form of PCNA has a pI of 6.96**
- **The modified form of PCNA has a pI of 4.76**
- **The modification to PCNA is apparently lost upon 2D PAGE and its subsequent electroelution from the gel**
- **The two forms of PCNA can be purified by ion-exchange chromatography taking into account their pIs**
- **The modified form of PCNA forms aggregates in solution upon purification**
- **Co-immune precipitation of PCNA with anti-p21 antibody produces a species of PCNA with an apparent pI different than that of the unmodified and modified forms**
- **GST-p21 pull down assays also produce a species of PCNA with a different pI identical to that of the co-immune precipitation**
- **Labeled recombinant GST-p21 is unable to bind PCNA in the Far Western experiments, which could be a function of the loss of modification mentioned above**
- **Sequence data has been generated from *in situ* digestion of PCNA resolved by 2D-PAGE followed by MS/MS using a ion trap mass spectrometer**
- **Although PCNA has been identified by mass spectrometry, the structure and location of the modification has yet to be elucidated**
- **The modified form of PCNA interacts with p21 with an apparent K_d of 1.07×10^{-10}**
- **The non-modified form of PCNA interacts with p21 with an apparent K_d of 1.04×10^{-5}**

REFERENCES

Acharya, S., Wilson, T., Gradia,S., Kane, M.F., Guerrette, S., Marsischky, G., Kolodner, R., and Fishel, R. (1996) hMSH2 forms specific mispair binding complexes with hMSH3 and hMSH6. Pro. Nat. Aca. Sci. 93:13629-13634.

Almendral, J.m., Huebsch, D., Blundell, P.A., Macdonald-Bravo, H., Bravo, R. (1987) Cloning and Sequence of the Human Nuclear Protein Cyclin: Homology with DNA Binding Proteins. P.N.A.S. 84: 1575-1579.

Applegren,N., Hickey, R.J., Kleinschmidt, A.M., Zhou, Q., Wills, P., Coll, J., Bachur, N., Swaby, R., Wei, Y., Quan, J.Y.,Lee,M.Y., and Malkas, L.H. (1995): Further characterization of the human cell multiprotein DNA Replication complex. J. of Cell. Biochem. 54 :32-46.

Boulikas, Teni. (1991) Reation between carcinogenesis, chromatin structure and poly(ADP) ribosylation. Anticancer Res.11:489-528.

Brenot-Bose, F., Gupta, S., Margolis, R.L., Fotedar, R. Changes in the subcellular localization of replication initiation proteins and cell cycle proteins during G1-to S-phase transition in mammalian cells. Chromosoma 103:517-527.

Brugarolas, J., Chandrasekara, C., Goron, J., Beach, D., Jacks, T., Hannon, G. (1995) Radiation-induced cell cycle arrest comprimised by p21 deficiency. Nature 377:552-557.

Chen, I-T., Smith, M., O'Connor, P., Fornace, A. (1995) Direct interaction of GADD45 with PCNA and evidence for competitive interaction of GADD45 and p21Waf1/Cip1 with PCNA. Oncogene 11(12):1931-1937.

Chen, I-T. Akamatsu, M., Smith, M., Lung, F-D., Duba, D., Roller, P.P., Fornace, A.J., O'Connor, P. (1995) Characterization of p21Cip1/Waf1 peptide domains required for cyclin E/Cdk2 and PCNA interaction. Oncogene 12(3):595-607.

Chen, J., Jackson, P., Kischer, M., Dutta, A. (1995) Separate domains of p21 involved in the inhibition of Cdk kinase and PCNA. Nature 374:386-388.

Chen, J., Chen, S., Saha, P., Dutta, A. (1996) p21Cip1/Waf1 disrupts the recruitment of human Fen1 by proliferating -cell nuclear antigen into the replication complex. P.N.A.S. 93:11597-11602.

Chen, J., Peters, R., Saha, P., Lee, P., Theodoras, A., Pagano, M., Wagner, G., Dutta, A. (1996) A 39 amino acid fragment of the cell cycle regulator p21 is sufficient to bind PCNA and partially inhibit DNA replication in vivo. Nuc. Acids Res. 24(9):1727-1733.

Chen, X., Zuo, S., Kelman, Z., O'Donnell, M., Hurwitz, J., Goodman,M.F. (2000) Fidelity of Eucayotic DNA Polymerase d Holoenzyme from *Schizocaccharomyces pombe*. JBC 275(23): 17677-17682.

Chen, Y.Q., Ciprano, S.C., Arenkiel, J.M., Miller, F.R. (1995) Tumor suppression by p21WAF1. Cancer Res. 55:4536-4539.

Chuang, L. S.-H., Ian, H.-I., Koh, T.-W., Ng, H.-H., Xu, G., Li, B.F.L. (1997): Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science 277: 1996-2000.

Coll, J., Weeks, J., Hickey, R., Schnaper, L., Yue, W., Brodie, A., and Malkas, L.H. (1996): The human breast cell DNA synthesome: Its purification from tumor tissue and cell culture. Oncol. Res. 8: 435-447.

Coll, J.M., Hickey, R.J., Cronkey, E.A., Jiang, H.-Y., Schnaper, L., Lee, M.Y.W.T., Uitto,L., Syvaoja,J.E., Malkas, L.H. (1997): Maping specific protein-protein interactions within the core component of the breast cell DNA synthesome. Onc. Res. 9:629-639.

Del Sal, G., Murphy, M., Ruaro, E.M., Lararevic, D., Levine, A.J., Schneider, C. (1996) Cyclin D1 and p21/waf1 are both involved in p53 growth suppression. Oncogene 12(1):177-185. Koeffler, H.P.(1995): Role of the Cyclin-Dependent Kinase Inhibitors in the Development of Cancer. Blood 86, 841-854.

Di Cunto, F., Topley, G., Calautti, Hsiao, J., Ong, L., Seth, P.K., Dotto, G.P. (1998): Inhibitory function of p21Cip1/WAF1 in differentiation of primary mouse keratinocytes independent of cell cycle control. Science 280:1069-1072.

Dimitrova, D., Gilbert, D.M. (2000) Tempoally coordinated assembly and disassembly of replication factories in the absence of DNA synthesis. Nature Cell Biol. 2: 686-694.

Eki, T. (1994):Poly(ADP-ribose) polymerase inhibits DNA replication by human replicative DNA polymerase alpha, delta, and epsilon *in vitro*. FEBS Letters 336:261-266.

Fotedar, A., Cannella, D., Fitzgerald, P., Rousselle, T., Gupta, S., Doree, M., Fotedar, R. (1996): Role for cyclin A-dependent kinase in DNA replication in human S phase cell extracts. JBC 271(49):31627-31637.

Fukuda, K., Morioka, H., Imajou, S., Ikeda, S., Ohtsuka, E., Tsurimoto, T. (1995) Structure-Function Relationship of the Eukayotic DNA Replication Factor, Proliferating Cell Nuclear Antigen. JBC 270(38): 22527-22534.

Gadbois, D., Peterson, S., Bradbury, E.M., Lehnert, B. (1995) CDK4/cylinD/PCNA complexes during Staurosporine-induced G1 arrest and G0 arrest of human fibroblasts. Exp. Cell Res. 220:220-225.

Gartel, A.L., Serfas, M.S., Tyner, A.L. (1996) p21-Negative regulator of the cell cycle. J. of Exp. Biol. and Med. 213:138-149.

Gomes, X.V., Burgers, P.M., Two Modes of FEN1 Binding to PCNA Regulated by DNA. (2000) EMBO 19(14):3811-3821.

Goubin, F., Ducommun, B. (1995) Identification of binding domains on the p21Cip1 cyclin-dependent kinase inhibitor. Oncogene 10(12):2281-2287.

Gorospe, M., Holbrook, N. (1996) Role of p21 in prosaglandin A2-mediated cellular arrest and death. Cancer Res. 56:475-479.

Gulbis, J., Kelman, Z., Hurwitz, J., O'Donnell,M., and Kuriyan, J. (1996): Structure of the
C-terminal region of p21 complexed with human PCNA. Cell 87:297-306

Haper, W., Adami, G., Wei, N., Keyomarsi, K., Elledge, S. (1993) The p21 cdk-interacting protein cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell 75:805-816.

Hedley, D.W., Rugg, C.A., and Gelberg, R.D. (1987): Association of DNA index and S-phase fraction with prognosis of node early positive breast cancer. Cancer Res. 47:4729-4735.

Hindges, R., Hubscher, U. (1997): DNA polymerase δ, an essential enzyme for DNA transactions. Biol. Chem. 378: 345-362.

Hirama, T., Koeffler, H.P.(1995) Role of the cyclin-dependent kinase inhibitors in the development of cancer. Blood 86 (3):841-854.

Hughli, Tony. (1989):Techniques in Protein Chemistry. New York, Academic Press.

Jackson, P.K., Chevalier, S., Phlippe, M., Kirschner, M.W. (1995): Early events in DNA replication require cyclin E and are blocked by p21CIP1. J. of Cell Biol. 130(4): 755-769.

Jonsson, Z.O., Podust, V.N., Podust, L.M., Hubscher, U. (1995) Tyrosine 114 is essential for the Trimeric Structure and the Functional Activities of Human Proliferating Cell Nuclear Antigen. EMBO 14(22): 5745-5751.

Kallioniemi, O.-P., Hietanen, T., Mattila, J., Lehtinen, M., Lauslahti, K., and Koivula, T. (1987):Aneuploid DNA content and high s-phase fraction of tumor cells are related to

poor prognosis in patients with primary breast cancer. Europ. J. cancer Clin. Oncol. 23:277-283

Kallioniemi, O,-P, Blanco, G., Alavaikko, M. Hietanen, T., Matilla, J., Lauslahti, K., Lehtinen, M., and Koivula, T. (1988): Improving the prognosis value of DNA flowcytometry in breast cancer by combining DNA index and S-phase fraction. A proposed classification of DNA histograms in breast cancer. Cancer 62:2183-2190.

Klinttenberg, S., Stal, O., Nordenskjold, B., Wallgren, A., Arvidsson, S., and Skoog, L.(1987):Proliferative index , cytosol estrogen receptor, and auxillary node status as prognostic predictors in human mammary carcinoma. Breast Canceer Res.Treat. 7: 99-106

Knibleiehler, M., Goubin, F., Escalas, N., Jonsson, Z., Mazarguil, H., Hubscher, U., Ducommun, B. (1996) Interaction studues between the p21Cip1/Waf1 cyclin-dependent kinase inhibitor and proliferating cell nuclear antigen (PCNA) by surface plasmon resonance. FEBS 391:66-70.

Kornberg, A., and Baker,T. (1992) in DNA replication , second edition New York, W.H. Freeman and Co.

Koundrioukoff, S., Jonsson, Z.O., Hasan, S., deJong, R.N., van der Vliet, P.O., Hottiger, M.O., Hubscher, U. A Direct Interaction between Proliferating Cell Nuclear Antigen (PCNA) and Cdk2 Targets PCNA-interacting Proteins for Phosphorylation. JBC 275(30):22882-22887.

Krishna, T., Xiang-Peng,K., Gary S., Burgers, G., Kuriyan, J. (1994): Crystal structure of the eukaryotic polymerase processivity factor PCNA. Cell: 79:1233-1243.

Lebel, M., Spillare E.A., Harris, C.C., Leder, P. (1999) The Werner Syndrom Gene Product Co-purifies with the DNA Replication Complex and Interacts with PCNA and Topoisomerase I. JBC 274(53): 37795-37799.

Levin D.S., Bai, W., Yao,N., O'Donnell, M., Tomkinson, A.E. (1997): An interaction between DNA ligase I and proliferating cell nuclear antigen: Implications for Okazaki fragment synthesis and joining. P.N.A.S. 94: 12863-12868.

Levin, D.S>, McKenna, A.E., Motycka,T.A., Matsumoto, Y., Tomkinson, A.E. (2000) Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. Current Biology 10:919-922.

Li , JJ, Kelly TJ. (1984) Simian virus 40 DNA replication *in vitro* Pro. Nat. Aca. Sci. 81:6973-6977.

Li, R., Waga, S., Hannon, G., Beach,D., Stillman, B. (1994) Differential effects by the p21 CDK inhibitor on PCNA-dependent replication and repair. Nature 371:534-537.

Li, R., Hannon, G.J., Beach, D., Stillman, B. (1996): Subcellular distribution of p21 and PCNA in normal and repair-deficient cells following DNA damage. Current Bio. 6 (2):189-199.

Lin, S., Hickey, R.J., Malkas, L.H. (1997): The isolation of a DNA synthesome from human leukemia cells. Leuk. Res. 21 (6):501-512.

Lin, S., Hickey, R.J., Malkas, L.H. (1997): The biochemical status of the DNA synthesome can distinguish between permenant and temporary growth arrest. Cell Gro. and Diff. 8:1359-1369.

Loor, G., Zhang, S-J., Zhang, P., Toomey, N.L., Lee, M.Y.W.T. (1997): Identification of DNA replication and cell cycle proteins that interact with PCNA. Nuc. Ac. Res. 25(24): 5041-5046.

Luo, Y., Hurwitz, J., Massague, J. (1995) Cell cycle inhibition by independent CDK and PCNA binding domains in p21Cip1. Nature 375:159-161.

Maga, G., Stucki, M., Spardari, S., Hubscher, U. (2000) DNA Polymerse Switching: I. Replication Factor C Displaces DNA Polymerase α Prior to PCNA Loading. JMB 295:791-801.

Malapetsa, A., Noe, A.J., Poirier, G.G., Desnoyers, S., Berger, N.A., Panasci, L.C. (1996): Identification of a 116 kDa protein able to bind 1,3-bis(2-chlorethyl)-1-nitrosourea-damaged DNA as poly(ADP-ribose) polymerase. Mut. Res. 363:41-50.

Malkas,L.H., Hickey, R. J., Li, C-J., Pederson, N., and Baril, E.F. (1990): A 21S enzyme complex from HeLa cells that functions in simian virus 40 DNA replication *in vitro.* Biochemistry 29:6362-6374.

Mattock, H., Jares, P., Zheleva, D.I., Lane, D.P., Warbrick, E., Blow, J.J. (2001) Use of Peptides from p21 (Waf1/Cip1) to Investigate PCNA Function in Xenopus Egg Extracts. Exp. Cell Res. 265: 242-251.

Mattock, H., Lane, D.P., Warbrick, E. (2001) Inhibition of Cell Proliferation by the PCNA-Binding Region of p21 Expressed as a GFP Miniprotein. Exp. Cell Res. 265: 234-241.

McDonald, R., Wu, G., Waldman, T., El-Deiry, S. (1996) Repair defect in p21WAF1/CIP1-/- human cancer cells. Cancer Res. 56:2250-2255.

Meyer, J.S.,Prey, M.U., Babcock, D.S. and McDivitt, R.W. (1986):Breast carcinoma, cell kinetics, morphology, stage and host characteristics. A thymidine lab study. Lab Invest.54:41-51.

Mossi, R., Keller, R.C., Ferrari, E., Hubscher, U. (2000) DNA Polymerase Swithing II: Replication Factor C Abrogates Primer Synthesis by DNA Polymerase α at a critical length. JMB 295:803-814.

Mozzherin, D.J., McConnell, M., Jasko, M.V., Krayevsky, A.A., Tan, C-K., Downey, K.M., Fischer, P.A. (1996); Proliferating cell nuclear antigen promotes misincorporaations catalysed by calf thymus DNA polymerase δ. 271(49): 31711-31717.

Mozzherin, D.J., Shibutani, S., Tan, C-K., Downey, K.M., Fischer, P.A. (1997): proliferating cell nuclear antigen promotes DNA synthesis past template lesions by mammalian DNA polymerase δ. P.N.A.S. 94: 6126-6131.

Nakanishi, M., Robetorye, R., Pereire-Smith, O., Smith, J. (1995) The C-terminal region of p21SDI1/Waf1/Cip1 is involved in proliferating cell nuclear antigen binding but does not appear to be required for growth inhibition. J. of Biol. Chem. 270(29):17060-17063.

Nakanishi, M., Adami,G., Robetorye, R., Noda, A., Venable,, S., Dimitrov, D., Pereira-Smith, O., Smith, J. (1995) Exit from G0 and entry into the cell cycle of cells expressing p21SDI1 antisense RNA. P.N.A.S. 92:4352-4356.

Nie, J., Sakamoto, S., Song, D., Qu, Z., Ota, K., Taniguchi, T. (1998): interactions of Oct-1 and automodification domain of poly(ADP-ribose) synthetase. FEBS Letters 424:27-32.

Pan, Z-Q., Chen, M., Hurwitz, J. (1993): The subunits of activator 1 (replication factor C) carry out multiple functions essential for proliferating-cell nuclear antigen-dependent DNA synthesis. P.N.A.S. 90:6-10.

Podust, V.O., Tiwari, N., Stephan, S., Fanning, E. (1998) Replication Factor C Disengages from Proliferating Cell Nuclear Antigen (PCNA) upon sliding Clamp Formation, and PCNA Itself Tethers DNA Plymerase δ to DNA. JBC 273(48):31992-31999.

Podust, V., Podust,L., Goubin F., Ducommun, B. and Hubscher, U. (1995): Mechanisnms of inhibition of proliferating cell nuclear antigen-dependent synthesis by the cyclin-dependent kinase inhibitor p21. Biochemistry 34:8869-8875.

Sato,T., Akiyama,F., Sakamoto, G., Kasumi, F., and Nakamura, Y. (1991): Accumulation of genetic alterations and progression of primary breast cancer. Cancer Res. 51:5794-5799.

Scott, M.T., Morrice, N., Ball, K.L. (2000) Reversable Phosphorylation at the C-terminal Regulatory Domain of $p21^{WAF1/CIP1}$ Modulates Proliferating Cell Nuclear Antigen Binding. JBC 275(15): 11529-11537.

Smith, M.L., Chen, I-T., Zhan, Q., Bae, I., Chen, C-Y., Gilmer, T.M., Kastan, M.B., O'Connor, P.M., Fornace, A.J., Jr. (1994): interaction of the p53-regulated protein Gadd45 with proliferatin g cell nuclear antigen. Science 266: 1376-1379.

Stillman, B. (1996) Cell cycle control of DNA replication. Science 274:1659-1664.

Tan, C-K., Castillo, C., So, A., Downy, K. (1986):An auxiliary protein for DNA polymerase δ from fetal calf thymus. J. Biol. Chem. 261:12310-12316.

Tom, T.D., Malkas, L.H., Hickey, R.J. (1996): Identification of multiprotein complexes containing DNA replication factors by native immunoblotting of HeLa cell protein preparations with T-antigen-dependent SV40 DNA replication activity. J. of Cell. Biochem. 63:259-267.

Tournier, S., Leroy, D., Goubin, F., Ducommun, B., Hyams, J. (1996) Heterologous expression of the human cyclin-dependent kinase inhibitor p21Cip1 in the fission yeast *Schizosaccharomyces pombe* reveals a role for PCNA in the chk1+ cell cycle checkpoint pathway. Mol. Cell Biol. 7:651-662.

Towbin, H. Straehelin, and Gordon, J. (1979):Electrophoretic transfer of protein from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Nat. Aca. Sci. 76:4350-4355.

Ueda, K. and Hayaishi, O. (1985) ADP-Ribosylation. Ann. Rev. Biochem. 54:73-100.
Gretch D.R., et al. (1987) Anal. Bichem. 163:270.

Umar,A., Buermeyer, A., Simon, J., Thomas,D., Clark, A., Liskay,R., Kunkel, T. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. (1996) Cell 87:65-73.

Waga, S., Hannon, G., Beach, D., Stillman, B. (1994) The p21 inhibitor of cyclin-dependent kinases controls DNA replication by interaction with PCNA. Nature 369:574-578.

Wang, Z.-Q., Auer, B., Stingl, L., Berghammer, H., Haidacher, D., Schweiger, M., Wagner, E.F. (1995): Mice lacking ADPRT and poly(ADP-ribosyl)ation develop normally but are suseptable to skin disease. Genes and Dev. 9:509-520.

Warbrick, E., Lane, D.P., Glover, D.M., Cox,L.S. (1997): Homologous regions of Fen1 and p21Cip1 compete for binding to the same site on PCNA: a potential mechanism to co-ordinate DNA replication and repair. Oncogene 14: 2313-2321.

Warbrick, E. (1998): PCNA binding through a conserved motif. BioEssays 20:195-199.

Wu, Y., Hickey, R.J., Lawlor, K., Wills, P., Yu, F., Ozer, H., Starr,R.,Yan-Quan , J., Lee, M and Malkas, L.H.(1994):A 17S form of DNA polymerase from mouse cells mediates the
in vitro replication of polyomavirus DNA. J. of Biochem. 54:32-46.

Xiong,Y., Zhang, H., Beach, D. (1992) D type cyclins associate with multiple protein kinases and the DNA replication and repair factor PCNA. Cell 71:505-514.

Yan G., Ziff, E.B. (1995): NFG regulates the PC12 cell cycle machinery through specific inhibition of the Cdk kinases and induction of cyclin D1. J. of Neurosc. 15(9): 6200-6212.

Yan, H., Newport, J. (1995) An analysis of the regulation of DNA synthesis by cdk2, Cip1, and licensing factor. J. Cell. Biol. 129(1):1-15.

Zhang, P., Sun,Y., Hsu, H., Zhang, L., Zhang, Y., Lee, M.Y.W.T. (1998) The interdomain connector loop of human PCNA is involved in a direct interaction with human polymerase δ. JBC 273(2): 713-719.

Myszka, D.G. (1999) Improving Biosensor Analysis. J. of Mel. Ruognit. 12:279-284.

Curriculum Vitae

Derek J. Hoelz — Graduate Research Assistant

Kent State University Kent, OH	B.A.	1990-1995	Biology
University of Maryland at Baltimore Baltimore MD		1996-present	

Papers:

Hoelz, D.J., Bechtel, P., Hickey, R.J., Malkas, L.H. (2001): Purification of the Malignant Form of Proliferating Cell Nuclear Antigen from Breast Cancer cells. Manuscript in preparation.

D. Hoelz, R. J. Hickey, and L. H. Malkas, Prokaryotic DNA Replication. (2001) *Encyclopedia of Life Sciences,* Nature Publishing (in press).

R. J. Hickey, D. Hoelz and L. H. Malkas, Eukaryotic DNA Replication. (2001) *Encyclopedia of Life Sciences,* Nature Publishing (in press).

Abstracts:

Hoelz, D.J., Sekowski, J.W., Hickey, R.J., Malkas, L.H. Identification of Mismatch Repair Proteins in the Human Cell DNA Synthesome. Presented at Graduate Student Research Day, 4-97.

Sekowski, J.W., Hoelz, D.J., Hickey, R.J., Malkas, L.H. Altered Fidelity in Cancer Cells. Presented at Graduate Student Research Day, 4-97.

Sekowski, J.W., Hoelz, D.J., Hickey, R.J., Malkas, L.H. Altered Fidelity in Cancer Cells. Presented at annual AACR meeting 4-97.

Sekowski, J.W., Hoelz, D.J., Malkas, L.H., Lu, A-L., Schnaper, L., Hickey, R.J. Structural and Functional Alterations of the DNA Synthesome-associated DNA Repair Proteins in Breast Cell Malignancy. Presented at annual AACR meeting 3-98.

Hoelz, D.J., Bechtel, P., Croisetiere, L., Sekowski, J.W., Freund, R., Malkas, L.H., Hickey, R.J. Inhibition of the Human Cell DNA Synthesome Through the Interactions of p21 and PCNA. Presented at annual AACR meeting 3-98.

Hoelz, D.J., Han, S.H., Bechtel, P.E., Freund, R., Schaper, L., Hickey, R.J., and Malkas, L.H. The Differential Effects of p21 on DNA Polymerase δ and DNA Replication Through its Interactions with the Different Forms of PCNA. Presented at the annual AACR meeting 4-99.

Hoelz, D.J., Park, M., Dogruel, D., Bechtel, P. E., Sekowski, J. W., Xiang, H.Y., Hickey, R.J., Malkas, L.H. Analysis of a Malignant Cell's DNA replication Apparatus by Mass Spectrometry. Presented at the annual AACR meeting 4-00.

Hoelz, D.J., Bechtel, P., Freund, R., Park, M., Hickey, R.J., Malkas, L.H. (2001) Isolation and Characterization of the Non-malignant Form of PCNA from MCF7 Breast Cancer Cells. Scientific Proceeding of the 92nd Annual meeting of the American Association for Cancer Research. 42: 894.

Tomic, D., Hoelz, D.J., Wills, P., Hickey, R.J., Schnaper, L., Lankford, C., Malkas, L.H. (2001) Detection of the Cancer Specific Form of PCNA by Elisa Assay. Scientific

Proceeding of the 92nd Annual meeting of the American Association for Cancer Research. 42: 466.

Hoelz, D.J., Bechtel, P., Hickey, R.J., Malkas, L.H. (2001) Differential binding of p21WAF1 by an altered form of proliferating cell nuclear antigen present in breast cancer cells. 24th Annual San Antonio Breast Cancer Symposium.

Patents:

Hoelz, D.J., Malkas, L.H., Hickey, R.J. (2001) The Isolation of Modified and Non-modified Forms of Proliferating Cell Nuclear Antigen (PCNA) for Diagnostic/ Drug development.

Awards and Recognition

Department of Defense Breast Cancer Fellowship. DAMD99-1-9273. July 1, 1999- June 30, 2002. $66,000. The Regulatory Interactions of p21 and PCNA in Human Breast Cancer.

Region 7 team member at Junior Olympic Nationals for gymnastics, second place team-1990.

Top 25 at Junior Olympic Nationals-1990

Full athletic scholarship including room and board to Kent State University-1990-1995

Graduate Research Assistantship (1997-1999), Department of Pharmacology and Experimental Therapeutics.

Workshops

Envited participant in the Molecular biology and pathology of neoplasia workshop, Keystone, Co.-1998

Skills:

Protein chromatography (Ion exchange, hydrophobic, affinity, gel filtration) agarose gel electrophoresis, 2D-PAGE, native-PAGE, SDS-PAGE, Western blotting, electroelution, immunoprecipitation, co-immunoprecipitation, enzymatic assays, SV40 DNA replication assay, molecular cloning, recombinant protein expression and purification, MALDI-TOF mass spectrometry, ESI-tandem mass spectrometry, BIACORE analysis of protein interactions and kinetics.

ORD Disclosure No. ____-_____-_____
(For office use only)

UNIVERSITY OF MARYLAND, BALTIMORE
INVENTION REPORT FORM

NOTE: ALL SECTIONS OF THIS FORM MUST BE COMPLETED.
INCOMPLETE SUBMISSIONS WILL BE RETURNED.

Definition of Invention: An invention is a novel, non-obvious and useful process, machine, article of manufacture, composition of matter, or related improvement. The inventive process consists of two steps: conception (mental formulation of the complete means by which a desired result is achieved) and reduction to practice (physically constructing or carrying out the mental formulation and testing by appropriate means to demonstrate the invention achieves the desired results).

Definition of Inventor: An inventor is one who contributed to the conception of the invention as well as its reduction to practice. The following are not necessarily co-inventors: co-author, collaborator, co-investigator. In order to be co-inventors, the parties must work toward a common goal, producing an invention by their aggregate efforts. While they need not physically work together, it is necessary that they each work on the same subject matter, each making some contribution to both the inventive thought and to the final result.

1. Title of Invention: The Isolation of Modified and Non-modified Forms of Proliferating Cell Nuclear Antigen (PCNA) for Diagnostic/Drug Development

2. Inventor Information:

Inventor #1. Full Name (printed or typed): Derek J. Hoelz
Home Address: 225 Scott Street
City: Baltimore State: MD Zip: 21230 Telephone: (410)230-0355
Department: Pharmacology Campus Address: 525C HSF
Title: Graduate Research Assistant
Dept Telephone : (410)706-1798 Fax: (410)706-0032 Email address: dhoelz@som.umaryland.edu
Citizenship: US SSN: 220-13-9272
Signature: Date:

Inventor #2. Full Name (printed or typed): Linda H. Malkas
Home Address: 239 Kensington Parkway
City: Abingdon State: MD Zip: 21009 Telephone: (410)569-8490
Department: Pharmacology Campus Address: 525D HSF
Title: Professor
Dept Telephone : (410)706-2313 Fax: (410)706-0032 Email address: lmalkas@som.umaryland.edu
Citizenship: US SSN: 106-48-7550
Signature: Date:

Inventor #3. Full Name (printed or typed): Robert J. Hickey
Home Address: 239 Kensington Parkway
City: Abingdon State: MD Zip: 21009 Telephone: (410) 569-8490

Department: Pharmacology
Campus Address: 7-029 BRB
Title: Associate Professor
Dept Telephone : (410) 328-7499 Fax: (410)706-0032
Email address: rhickey@som.umaryland.edu
Citizenship: US
SSN: 127-40-4954
Signature:
Date:

[Use extra sheets if necessary]

3. Advisor approval for student submissions (if applicable):

Name and Title: Linda H. Malkas, Professor
Date: 1-11-01

4. Is this submission related to another invention disclosure previously submitted to ORD?

Yes ☒
No.☐

If yes, what is the title of that submission?

5. Description of Invention. (NOTE: Use additional sheets if necessary. You may also attach a manuscript, research proposal, drawing, or any other material to support the description below.)

A. In your own words, how would you briefly describe the invention?

This invention effectively purifies two different forms of the mammalian DNA replication protein proliferating cell nuclear antigen (PCNA). Analysis of malignant versus non-malignant cells grown in cell culture and from tissue has consistently shown the appearance of a modified and a non-modified form of PCNA. However, comparison of PCNA contained in the malignant cells with that contained in the non-malignant cells demonstrated an exclusive presence of a non-modified form of PCNA in the malignant cells. In other words, the malignant cells contain a modified and a non-modified form of PCNA, while the non-malignant cells contain a single form, the modified form. It is therefore reasonable to assume that this form of PCNA found in malignant cells (non-modified) may play a role in the promotion and/or maintenance of a malignancy. Isolation and purification of these two forms is absolutely essential for the development and testing of diagnostic methodologies for the detection and quantitation of the two forms of PCNA found in biological samples. In addition, isolation and purification of PCNA is essential for study of the biological processes involved in (a) generation of the two forms of PCNA, (b) the significance the two forms have on basic biological processes of the cell, (c) their differences and effects these differences have on non-malignant and malignant cellular function, and (d) the development of possible therapeutics to promote, stop, circumvent, and/or prevent the aforementioned processes.

Mammalian cells grown in culture or tissue specimens are first homogenized using a Dounce homogenizer. The cellular homogenate is subsequently centrifuged to separate the nuclei from the cytosol. Microsomes and mitochondria are then pelleted out of the cytosolic fraction by high-speed centrifugation and proteins present in the nuclei are extracted using a buffer containing 150mM potassium chloride. The insoluble nuclear membrane and DNA of the nuclear fraction are pelleted by high-speed centrifugation and the clarified cytosolic and nuclear fractions are combined. PCNA is purified from other proteins present in the cytosolic/nuclear fraction by passage over a strong anion exhange column. The flow through fraction of the anion exchange column contains PCNA, which is further purified by hydrophobic chromatography. These chromatographic steps produce a fraction enriched for the two forms of PCNA (cancer-specific and non-malignant). Separation of the two forms is then accomplished by cation-exchange chromatography using a selective pH that promotes the binding of the non-malignant form of PCNA while preventing the binding of the cancer-specific form of PCNA. The two isolated forms are finally purified to essential homogeneity by anion-exchange chromatography.

B. For marketing purposes and federal reporting requirements, please provide a list of five keywords

Cancer , Protein purification, PCNA, post-translational modification, DNA synthesome

6. Ownership and Results Demonstrating the Concept is Valid.

A. When was the invention first thought of? Is this date documented? Where?

This invention was first thought of on June 6, 2000 and the date is documented in the PCNA purification laboratory notebook number 1 of Derek J. Hoelz.

Who employed each co-inventor at that time?

#1. University of Maryland, Baltimore #2.

#3. #4.

B. Has the invention been tested experimentally? Please state if you have preliminary results, animal or laboratory models, prototypes, or clinical tests. (Simply reference appropriate sections if you have already included this information. You may also attach additional pages if necessary.)

The invention has been tested experimentally. The cancer-specific and non-malignant forms of PCNA separated by the above methods have been tested for purity using slot blot, two-dimensional polyacrylamide gel electrophoresis, and Western blot experiments.

If different from above, who has employed each co-inventor during the period of experimental testing?

#1. #2.

#3. #4.

C. Have any of the materials used in the conception and reduction to practice of this invention been acquired from an outside source, for example under a Material Transfer Agreement (excluding outright purchases)?

Yes ☐ No.☒ If yes, please provide details of the transaction below.

Material: Source:

Explanatory Notes:

Identify Agreement:

D. Who provided external support for the work that led to the invention? External support includes funding by a nonprofit or for-profit entity of: This work was supported by grants issued to Linda H. Malkas and Robert J. Hickey from the National Institutes of Health and a grant issued to Derek J. Hoelz from the U.S. Army.
(A) an entire research project;
(B) any portion of a co-inventor's salary or the salary of individuals working under a coinventor's supervision on a research project; or
(C) supplies, reagents, animals, tissues, cells, or any other materials used in a research project.

1.Entity: Cells, columns, consumables	Grant # (if applicable): CA83199 to LHM
Termination date: 2005	Brief explanatory notes:
2.Entity: Cells, columns, consumables	Grant # (if applicable): #CA74904 to RJH
Termination date: 8/00	Brief explanatory notes:
3.Entity: Cells, columns, consumables	Grant # (if applicable): CA73060 to LHM
Termination date: 2000	Brief explanatory notes:
4.Entity: Salary , fringe benefits, student fees to DJH	Grant # (if applicable):
Termination date: 2002	Brief explanatory notes:

7. Uses and Applications. What are the possible uses for the invention? What products could be developed? In addition to immediate applications, are there other uses that might be realized in the future?
The possible uses for this invention include development of a diagnostic assay to test for malignancies of the breast and other tissues and its use as a standard for comparing stages of malignancies detected using this diagnostic assay. The invention would also be useful for determination of protein structure for the design/screening of therapeutics targeted to either of the two forms of PCNA.
8. Novel Features. Pick out and expand on the novel and unusual features of the invention. How does it differ from present technology? What problems does it solve? What advantages does it possess?
The novel feature of this invention is that it successfully purifies two nearly identical species of the same protein by utilizing the differences in charge (pI). The invention solves the problems encountered when analyzing the differences in the two forms of PCNA. It can serve as an assay standard for diagnostic tools utilizing the two forms of PCNA and as an essentially pure target for the analysis and development of new anti-neoplastic agents.
9. Obstacles. Does the invention have any disadvantages or limitations? How can they be overcome?
The nmPCNA purified using these methodology forms aggregates and becomes insoluble when the sample is freeze-thawed. The result is a loss of protein upon storage.

10. Future Research Plans
A. What additional research is needed to complete development and testing of the invention?
Proper buffers and conditions must be developed for stabilization of purified nmPCNA enabling long-term storage and conditions for shipment. The purified nmPCNA must be tested to demonstrate that the protein has the correct pI.
B. Is this research presently being undertaken? Yes☒ No☐
If yes, under whose sponsorship? NIH and US Army (see 6.B.)

Entity: R01 awarded to LHM and pre-doctoral fellowship awarded to DJH.	Grant # (if applicable): CA83199
Termination date: 2005 and 2002	Brief explanatory notes:

11. Inventors' Publications/ Presentations.
A. Has the invention or a similar invention in whole or in part been described in a publication? ("Publication" for this purpose includes abstracts of talks, news articles, scientific papers, poster sessions,

web postings, thesis, etc.)
Yes ☒ No ☐

PLEASE PROVIDE EXACT DETAILS INCLUDING DATES AND ATTACH COPIES OF ANY PUBLICATIONS TO THIS SUBMISSION;

Date: 4/2001 Citation: Isolation and Characterization of the Non-Malignant Form o PCNA from MCF7 Breast Cancer Cells.

Brief Explanatory Notes: Poster presented at the annual meeting of the American Association of Cancer Research

Date: Citation:

Brief Explanatory Notes:

Date: Citation:

Brief Explanatory Notes:

B. Has the invention been described orally at meetings or seminars?

Date: Citation:

Brief Explanatory Notes:

Date: Citation:

Brief Explanatory Notes:

Date: Citation:

Brief Explanatory Notes:

C. Are you planning any disclosure of the invention in the future? What type of disclosure and approximately when?

Date: Citation:

Brief Explanatory Notes:

Date: Citation:

Brief Explanatory Notes:

Date: Citation:

Brief Explanatory Notes:

PLEASE NOTE: In order to complete review of your invention report and timely filing of a patent application, if appropriate, ORD may require up to 60 days before any future publication date.

12. Prior Art

A. A literature search should be done by the inventor to identify publications relevant to this invention. Please list publications and any related patents you may know of. Use additional sheets if necessary. Free patent searches can be performed on-line using the IBM searcher (http://www.patents.ibm.com/) or the USPTO home page (http://www.uspto.gov/patft/index.html). Free Medline literature searches can be performed using the PubMed searcher (http://www.ncbi.nlm.nih.gov/PubMed/). Please attach a print-out of relevant hits.

B. Do you know of relevant information presented at a public talk, trade fair, sales catalogue? Has the invention or a similar product been used publicly or has it been offered for sale?

13. Marketing Information

A. Has there been any commercial interest in this invention? Please name companies and specific contacts.
Minerva Pharmaceuticals. Ken Blackman, Gary Siebert

B. What companies do you think may be interested in this invention? Why?
Minerva Pharmaceuticals would be interested in this purification scheme because they are currently attempting to develop a diagnostic test utilizing the two forms of PCNA.

C. What companies make products currently in the market that compete with your invention?
None.

D. What is the potential size of the market? For example, if the invention is a new therapeutic agent, can you give an estimate of the number of people afflicted in the U.S. and abroad?
>200,000/year in the United States

E. If your invention would be primarily used in countries other than the U.S., please suggest which countries and give a brief explanation.

F. What other research groups are working on similar inventions and where are they located?
None.

Please note: If ORD pursues patent protection based on this invention report, it will be necessary to have a commercial licensee identified within approximately 18 months of patent application filing.

ORD will contact you to schedule a meeting with our patent attorney and technology administrator following receipt of this report.

THANK YOU

Title: DNA Replication of Prokaryotes and Yeast

Authors: Derek J. Hoelz, Robert J. Hickey, Linda H. Malkas

Affiliation: University of Maryland School of Medicine
Department of Pharmacology and the Greenebaum Cancer Center
Health Sciences Facility I
685 W. Baltimore Street
Baltimore, MD 21201

Phone: 410-706-2313

Fax: 410-706-0032

Email: lmalkas@som.umaryland.edu

I. *ESCHERICHIA COLI*: THE ARCHETYPE MODEL FOR REPLICATION OF A DNA GENOME

IA. General Scheme for *Escherichia coli* (*E. coli*) DNA Replication

Over the last 30 years an abundance of biochemical and genetic data have accumulated identifying at least 30 proteins required to replicate the bacterial *E. coli* DNA genome (1). The *E. coli* genome consists of a relatively small circular chromosome comprised of 4.7 million base pairs (Mbp), which, upon receiving the proper cellular signals, is replicated prior to cell division. Chromosomal DNA replication in *E. coli* can be divided into three discrete stages beginning with initiation. Once initiation occurs, a 42 minute constant period (C period) is required for chromosomal replication. After initiation, the elongation stage takes place. The elongation stage is the most time consuming phase of the replication process and is where the vast majority of the DNA is synthesized. The final stage of *E. coli* DNA replication, designated the termination phase, marks the end of the C period. It is accompanied by the generation of two independent daughter chromosomes. The cells then enter cell division (D period) which lasts approximately 20 minutes yielding two genetically identical bacteria. Because of the 20 to 60 minute doubling time and the relatively small DNA content of the chromosome, *E. coli* was the first organism used to study the mechanisms of DNA replication. In this section the specific mechanisms involved in *E. coli* DNA replication will be presented.

I B. The Initiation of *E. coli* DNA Replication

Chromosomal DNA replication in *E. coli* is initiated at a unique sequence, the chromosomal origin of replication or *oriC* (2). *OriC* is composed of a minimal region of 245 bp that contains three 13 bp AT-rich domains and four 9 bp sequences (dnaA boxes) that bind the initiation protein dnaA, Figure 1. Initiation at *oriC* occurs through the binding of the dnaA protein to the 9-bp repeats in the presence of ATP and another protein, HU, Figure 2. Twenty to 40 dnaA monomers organize with HU into a protein core to which the *oriC* DNA is wrapped, forming the initial complex. Formation of the initial complex places the three 13 bp AT-rich domains in close proximity to the dnaA core, and in an ATP-dependent manner, dnaA melts the AT-rich domains sequentially

starting with the domain closest to the core. The melting of the AT-rich domain results in the formation of a bubble of single-stranded DNA that quickly becomes coated with single-stranded DNA-binding protein (SSB). Binding of SSB to single-stranded DNA in effect stabilizes the single-stranded structure. A dimeric protein complex dnaB-dnaC is then directed towards the single-stranded DNA through an interaction with dnaA. Here, dnaB recognizes single-stranded DNA and using its intrinsic 5'→3' helicase activity further unwinds the DNA in an ATP-dependent fashion. Extension of the unwound region of the DNA by dnaB creates a prepriming complex. At this point, the DNA is ready for elongation.

I C. *E. coli* DNA Replication Elongation Events

Once the prepriming complex has formed, single-stranded DNA is accessible, allowing for the synthesis of a new daughter strand of DNA on the parental template. However, in order for the replicative polymerase, DNA polymerase III holoenzyme (DNA pol III HE), to synthesize DNA, it must have a 3'-OH terminus from which it will elongate. This function is accomplished by a primosome. Although there is much confusion about what exactly constitutes a replicative primosome in *E. coli*, two proteins, the helicase dnaB, and an RNA primase, dnaG, are absolutely essential. However, for the purpose of this article, we will concentrate on replication initiated from *oriC*, and thus the primosome to which we will refer simply consists of dnaB, dnaC, and dnaG.

DnaG, an RNA primase, interacts with dnaB, the helicase recruited to the prepriming complex stimulating the activity of dnaB. In turn, this interaction stimulates the primase activity of dnaG, which then synthesizes a 10-12 ribonucleotide (RNA) primer on both strands of the DNA duplex at the replication fork. The DNA pol III HE, the replicative DNA polymerase, is now directed to the replication fork by an interaction with the dnaB-dnaG complex, Figure 3. At the fork, DNA pol III HE begins synthesizing nascent strands of DNA from the primers.

DNA pol III HE is a protein complex composed of at least 10 different proteins and can carry out DNA synthesis at a rate of nearly 1000 nucleotides/minute. As the replication

fork progresses, DNA pol III HE must synthesize DNA onto two separate templates. As stated above, the DNA strand oriented in the 5'→3' direction, the same direction as replication fork movement, is referred to as the leading strand, while the template orientated in the 3'→5' direction is considered the lagging strand. Leading strand synthesis is accomplished by DNA pol III HE synthesizing extensive lengths (>0.5 Mbp) of DNA. However, due to the antiparallel nature of the DNA and the ability of DNA pol III to synthesize DNA exclusively in the 5'→3' direction, a problem occurs when attempting to synthesize DNA on the lagging strand. How do you synthesize DNA in a 5'→3' direction while moving in a 3'→5' direction? Okazaki and co-workers first proposed the solution to this problem in the 1970s, and the answer lay in the discontinuous synthesis of small 1000-2000 bp DNA fragments subsequently termed Okazaki fragments (3). Synthesis of the lagging strand is completed by looping small sections of the DNA strand and placing them in the 5'→3' orientation. DNA pol III HE is then able to synthesize DNA onto these small looped sections of DNA, creating an Okazaki fragment. Once complete, the DNA pol III HE cycles to a new looped segment of the DNA strand 5' to the recently completed Okazaki fragment and again begins to synthesize DNA. The end result is a long stretch of Okazaki fragments on the lagging DNA strand that can be joined or ligated together to form one long contiguous stretch of DNA.

DNA pol III HE can be resolved into its individual protein components. Isolation of these protein components has afforded valuable data about the enzymology and stoichiometry of DNA pol III HE. The core of DNA pol III HE is composed of three proteins, α, ε, and θ. The α subunit contains the catalytic activity and is able to synthesize DNA in the 5'→3' direction, while the ε subunit contains a 3'→5' exonuclease and serves a structural function. The contribution(s) of θ to the core has yet to be determined. Although the pol III core has catalytic activity, its DNA synthetic processivity is extremely low. In other words, the pol III core is able to only synthesize very short stretches of DNA using a parental template. This processivity is increased by the addition of τ. τ causes two pol III core complexes to dimerize, forming a DNA pol III' complex. Through an interaction with dnaB, τ is able to coordinate the helicase

activity of the replication fork with leading strand synthesis. Addition of another group of proteins, the γ complex, to the DNA pol III' complex yields yet a more processive complex, DNA pol III*.

The γ complex consists of six proteins, including τ. A τ dimer interacts with a γ dimer, anchoring the γ complex to DNA pol III'. τ and γ are protein products of the same gene, dnaX; however, γ is truncated by a translational frame shift and is unable to dimerize the DNA pol III cores or interact with dnaB. τ and γ are now able to associate with two other proteins δ and δ'. High-affinity binding of δδ' to τγ requires a preceding interaction with another two proteins, χ and ψ (4,5). ψ, an insoluble protein by itself, tightly binds χ and, although not a lot is known about the protein, appears to form a bridge between χ and the γ complex through its interaction with δδ'. χ functions by an interaction with dnaG primase and SSB, and it has been proposed that through this interaction the "hand-off" of the primer from dnaG to the the DNA polymerase is accomplished during Okazaki fragment synthesis of the lagging strand. χψ is now bound to δδ', which, in turn, is bound to τγ completing the functional γ complex. Finally, the addition of a β protein dimer converts the DNA pol III* complex to the highly processive DNA pol III HE (6,7).

The β subunit of DNA pol III HE is a member of a group of proteins called processivity factors. The β dimer encircles the DNA forming a "sliding clamp." The β sliding clamp then interacts with DNA pol III* through an interaction with the α subunit and, in effect, tethers the polymerase to the DNA, greatly increasing its processivity. In order for β to act as a processivity factor, it must first be loaded onto the DNA and positioned in such a way as to promote DNA polymerization. As it turns out, β clamp loading onto DNA is accomplished through the action of the γ complex of DNA pol III* mentioned earlier.

The γ clamp-loading complex has been an area of intense study. This is because the complex is not only responsible for clamp loading but also for positioning β at the primer terminus. Clamp loading is an ATP-dependent reaction requiring the hydrolysis of two molecules of ATP per clamp-loading event. The reaction is initiated at the replication

fork where a β dimer is bound to the δ subunit of the γ complex. The interaction with δ opens the dimer and loads the clamp onto the DNA in an ATP-independent manner. Recently, it has been suggested that δ may also unload β from the DNA following completion of the synthesis of an Okazaki fragment. Once loaded onto the DNA, β must be positioned at or near the 3'-OH terminus of the RNA primer so as to allow for an interaction with the α catalytic subunit of the DNA polymerase. The translocation and positioning of β on the DNA is what requires the hydrolysis of ATP, an activity intrinsic to the τγ subunits. With β now loaded and positioned at the primer termini of the leading and lagging strands, elongation commences.

The replication forks now begin to move bi-directionally from the origin. The replication "bubble" formed at the end of initiation begins to expand giving rise to a θ structure consisting of the parental chromosome and two newly replicated half circles of daughter chromosome. Elongation is completed approximately 40 minutes following the onset of initiation, and termination occurs at specific sequences on the *E. coli* chromosome approximately 180° away from *oriC*.

I D. *E. coli* DNA Replication Termination

Although the events involved in the termination of *E. coli* replication are poorly understood, some key features have been described that seem to be essential for the generation of two daughter chromosomes (8). At the end of elongation, the replication forks eventually meet at a section of the *E. coli* chromosome approximately 180° from *oriC,* Figure 4. It is in this part of the genome that the replication forks will encounter specific DNA sequences called termination (*ter*) sequences. The *ter* sequences are arranged on the chromosome in two groups of three direct repeats, each group having the opposite orientation. The *ter* sequences function through the binding of a monomer of the protein Tus. The crystal structure of Tus complexed to a *ter* sequence has given insight into its binding and function in termination. Within the protein itself, Tus contains β sheets that contact the major groove of the DNA and give Tus its binding specificity. Projected over the DNA binding domain is Tus' replication fork arresting

domain. Tus arrests the movement of the replication fork specifically by inhibiting the helicase activity of dnaB, and, although an interaction between dnaB and Tus has been demonstrated *in vitro*, the significance of this interaction *in vivo* has yet to be determined. Another interesting aspect of Tus is that its helicase-arresting or contra-helicase activity is polar. Imagine a replication fork moving clockwise along the circular *E. coli* chromosome. Upon encountering the first group of three *ter* sequences complexed with Tus monomers, the replication fork moves through the sequences unabated. Due to the orientation of the ter sequences, Tus is unable to block the helicase activity, and therefore is unable to arrest the replication fork. However, when the replication fork encounters the next group of three *ter* sequences with the opposite orientation, helicase activity is inhibited and replication fork movement is arrested. The arrest of replication fork movement, however, is only one of the steps involved in termination. Another step is the separation of the two daughter chromosomes. Through the action of topoisomerase IV, an enzyme that makes a cut in both strands of a DNA duplex, the daughter chromosomes are separated from one another.

II. BACTERIOPHAGE MODELS OF DNA REPLICATION

Bacteriophages (phages) have proven to be valuable model systems for the study of DNA replication. First, phage genomes are considerably smaller than bacterial genomes and easier to manipulate. Their small genome size makes them rely on many bacterially encoded proteins for their replication. Finally, phage replication is not restricted by the temporal events of cell division.

Several phage models have been developed and used to study the mechanisms of DNA replication in prokaryotes. Some of these models include single-stranded DNA phages represented by ΦX174, and others include double stranded DNA phages represented by T4, T7 and λ. Although the replication of the DNA of these phages contain many similarities to one another, each of the phage mechanisms is unique. For the purpose of this article, only a few of these mechanisms (i.e., for λ, T7, T4, and ΦX174) will be described. Each of these phage model systems has contributed tremendously to our understanding of the complex series of events involved in replicating DNA.

II A. λ DNA Replication

The λ phage is packaged with a double-stranded (ds) linear chromosome that, upon infection of a bacterium, becomes circular, or cyclizes, through the hybridization of two cohesive ends called *cos* sites, Figure 5. Formation of these circular molecules allows for replication to occur similar to that described for the circular *E. coli* chromosome (see above). Requiring only two phage-encoded proteins, protein O (λO) and protein P (λP), the λ chromosomal replication initiates from a specific sequence, *oriλ*. Unlike the dnaA dependent initiation of *E. coli* at *oriC,* initiation at *oriλ* requires the action of a cellular RNA polymerase (RNAP). RNAP synthesizes an RNA transcript from one of two promoters present in the origin, through the *oriλ* region, which is composed of four direct 18bp repeats and a 35-40bp AT-rich region. Synthesis of RNA through the *oriλ*, promotes the binding of λO to the 18bp sequences much like the binding of dnaA to *oriC* in *E. coli* unwinding the AT-rich region. Priming and replication occur with the recruiting of dnaB, the cellular helicase, through the interaction of λP with λO. DnaB further unwinds the DNA at the *oriλ* and localizes dnaG to the newly formed single-stranded DNA enabling it to synthesize the RNA primers on the leading and lagging strands. Like cellular replication, DNA pol III HE is then recruited to the replication fork leading to semiconservative DNA synthesis on the λ parental DNA template. When the replication forks on the λ DNA meet on the other side of the circular genome replication is terminated and two daughter molecules are separated. After a few rounds of this circle-to-circle replication, λ switches to a rolling-circle mechanism of replication. This requires the action of one additional phage-encoded protein, Gam. Rolling-circle replication proceeds when one strand of the double-stranded circular λ chromosome is cut, thus enabling the free 3' end to be used as a primer for DNA synthesis. Subsequent DNA synthesis produces a long stretch of double-stranded, tandomly repeated, linear λ chromosomes, called concatemers. The concatemers are eventually cut at the *cos* sequences and packaged into the phage heads.

II B. T7 DNA Replication

The T7 bacteriophage, like λ, contains a linear chromosome. However, in contrast to λ, the T7 viral DNA encodes nearly all of the proteins needed for its own DNA replication, and does not cyclize after infection. Replication of the linear T7 DNA template presents a major problem, Figure 6. As mentioned above, the DNA polymerases are able to synthesize exclusively in the 5'→3' direction and can only elongate from an existing 3'-OH primer terminus. As with *E. coli*, T7 also utilizes RNA primers, which presents a problem at the 5' ends of a linear template. After the RNA primer synthesized on the 5' end is degraded, the result is a single-stranded 3' DNA overhang, Figure 6C. Therefore, if these DNA overhangs are not corrected the linear template would get progressively smaller, eventually making the virus unviable. To correct 3' overhangs at the ends of its genome, phages T7 and T4 have developed clever DNA replication mechanisms, which will be discussed below.

T7 DNA replication initiates with the action of a virally encode RNA polymerase (vRNAP). Synthesis of RNA by the vRNAP from two promoters contained in the T7 origin, *oriT7*, primes the initiation reactions, and is therefore considered R-loop dependent, Figure 6A. *OriT7* consists of a 200bp region, which, in addition to the promoter elements, embodies a 61bp AT-rich region that contains T7 gene 4 product (g4P)-binding sites. The resultant RNA/DNA hybrid displaces the DNA strand at the AT-rich region, allowing for g4P to bind to the non-template strand. The helicase activity of g4P further unwinds *oriT7* in an ATP-dependent fashion. Translocating in the 5'→3' direction, g4P unwinds the DNA until the first priming site is reached. At the priming site, g4P then synthesizes an RNA primer with its intrinsic primase activity. Exchange of the RNA primer from g4P to the T7 DNA polymerase, g5P, establishes the first Okazaki fragment, which will then serve as a 3'-OH primer terminus for leading strand DNA synthesis and the progression of the replication fork. In the opposite direction, the RNA transcript initially synthesized in the initiation reaction serves as the primer for leading strand synthesis, and another replication fork is assembled and DNA synthesis proceeds bi-directionally. Termination of DNA synthesis occurs when the replication forks reach the ends of the linear template.

After termination, the T7 DNA now contains single-stranded 3' DNA overhangs. Resolution of the DNA loss associated with the single-stranded DNA overhangs is accomplished by hybridization of the complementary terminally repeated (TR) ends present on each adjacent newly replicated chromosome, Figure 6D. This hybridization of the newly replicated DNA forms extensive lengths of end-to-end chromosomes called concatemers. One T7 genome is subsequently packaged into one phage head by cutting the concatemer at specific sites called *pac* sites present between each genome, Figure 6E. After assembly of the phage particle is complete, the cell lyses and the viral particles are ready to infect another bacterium.

II C. T4 DNA Replication

Shortly after infection, the T4 phage takes over host DNA, RNA, and protein synthesis, kidnapping the cell's RNA and protein synthetic machinery to produce its viral proteins. T4 then degrades the bacterial genome increasing the free nucleotide pool to the advantage of its own DNA replication. This complete dominance over cellular activities makes the T4 phage an attractive model for the study of DNA replication, allowing researchers to monitor phage biosynthetic activities in the absence of any background activity of the host. In addition, the genome of T4 contains a double-stranded linear chromosome that, upon infection, does not cyclize. Encoding nearly all proteins needed to carry out DNA synthesis, the T4 phage has proven to be an extremely valuable model in the study of DNA replication.

Interestingly, the T4 mechanism of replication occurs in two stages. In stage 1, initiation of replication occurs through an R-loop dependent mechanism, Figure 7. The host's RNAP synthesizes RNA primers from multiple promoters located throughout the T4 genome, and replication is initiated at seven or more origins. Later in infection, T4 no longer initiates by an R-loop dependent mechanism. Instead, in stage 2, T4 initiates replication using an alternative mechanism, which is referred to as D-loop dependent, Figure 7 (9). Initially discovered in T4, D-loop dependent replication requires the formation of recombination intermediates for initiation to occur. Referred to as recombination dependent replication (RDR), study of T4 DNA replication initiated at D-

loops proved the long hypothesized involvement of the genetic recombination process in DNA replication.

As with the replication of the linear T7 template, the replication of the T4 genome from defined origins in stage 1 leads to DNA molecules that contain 3' single-stranded DNA overhangs at their terminal sequences. However, these single-stranded 3' DNA overhangs do not hybridize end-to-end like in T7, but are then able to invade and displace a homologous strand of DNA on an adjacent T4 chromosome forming recombination intermediates called D-loops, Figure 7. The 3'-OH of the invading strand then serves as a primer for leading strand DNA synthesis and an assembly site for the replication forks. D-loops form because the T4 genome contains redundant sequences in combination with the packaging of slightly more DNA than that contained in a single chromosome (see below). Currently, there are three models for DNA synthesis occurring at D-loops; unidirectional, bi-directional, and tri-directional models, Figure 8. However, the actual mechanism(s) that take place *in vivo* has yet to be determined. Unidirectional replication is similar to bi-directional replication in that both models resolve the cross-strand structure of the recombination intermediate by cleavage of a DNA strand. The models differ only in the strand of DNA that is cleaved. The third or tri-directional model, on the other hand, results from the total absence of DNA cleavage. After two replication forks are assembled at the D-loop, similar to the bi-directional model, the forks leave the area. The resultant Y-shaped structure could initiate a third replication fork leading to another round of replication. Regardless, the replication fork(s) initiated from D-loops will eventually reach the end of the DNA template, and 3' single-stranded DNA will be regenerated. The 3' single-stranded DNA will then be able to invade a homologous sequence forming another D-loop beginning the process over again. This self-regenerating mechanism ultimately results in the formation of long concatemers of tandemly repeated T4 genomic DNA much like those developed by the λ and T7 phages. The long concatemeric DNA is then packaged into T4 phage heads through a mechanism referred to as headful packaging. Headful packaging obtains its name because the T4 phage does not package a single genome into its head through cleavage at *cos* or *pac* sites as do λ and T7, respectively. Instead, T4 packages as much DNA into its head as space

will allow. Cleavage of the DNA takes place after the T4 phage head is full. This results in the packaging of slightly more DNA than that contained in a single T4 genome. The excess DNA contains a section of the next adjacent T4 genome redundant in sequence to the first, and these redundant sequences at the ends of the packaged DNA will serve as the invading strands for D-loop formation producing the long concatemers of the next infection cycle.

II D. ΦX174 DNA Replication

Unlike the phages that have been discussed so far, the ΦX group represent a new group of *E.coli* phages that contain a small single-stranded circular DNA chromosome. These small phages are well suited as models for the study of DNA replication for many reasons. For example, ΦX174's single-stranded chromosome has proven easy to isolate and prepare intact, and, also, its small size limits the number of virally encoded proteins, and therefore the phage must rely on many of the host's cellular proteins to direct its DNA replication. In addition, the ΦX174 life cycle has attributes that make it attractive for study. After infection of a bacterium, the single-stranded circular ΦX174 genome is injected into the cell. Once inside the bacterium, the single-stranded chromosome is converted to double-stranded DNA, which then can be replicated using a mechanism similar to *E. coli* chromosomal replication. Much of the research done on the ΦX174 phage concentrates on the conversion of single-stranded circular DNA to double-stranded DNA, and we will therefore concentrate on the reactions that take place in this stage.

The conversion of ΦX174 single-stranded DNA to double-stranded occurs in two phases. In phase 1, a prepriming intermediate is formed, which will be elongated by the DNA pol III HE (see section on *E.coli* DNA replication) in phase 2 (10). Phase 1 requires the action of a group of *E. coli* proteins that include the dnaB-dnaC complex that is involved in *oriC*-directed chromosomal replication in *E. coli*. In addition to the dnaB-dnaC complex, four additional proteins are involved in the formation of a prepriming complex. These proteins are: PriA, a DNA-dependent ATPase with translocase and 3'→5' helicase activities; PriB and PriC, whose function still remain elusive except for a redundant structural role; and dnaT, Figure 9. This group of proteins assembles at a specific site on

the ΦX174 single-stranded DNA template and is referred to as the primosome assembly site (PAS). Analysis of the PAS site demonstrated that the sequence of the site is not directly responsible for protein binding, but instead, primosome assembly is most likely dependent on the secondary DNA structure of the PAS site, Figure 14. In other words, primosome assembly results from topological recognition of structures formed by the interaction of the single-stranded DNA with itself.

Preprimosome assembly initiates with recognition of the PAS site by PriA. PriA binding is followed by the recruitment of PriB, PriC, dnaT, and dnaB, respectively; forming the preprimosome complex of phase 1. Formation of a functional primosome is completed with the association of dnaG, and marks the beginning of phase 2. DnaG of the primosome complex is then able to synthesize an RNA primer on the single-stranded DNA template followed by the initiation of nascent DNA synthesis by DNA polymerase I (pol I). After synthesis of about 300bp, DNA pol III HE replaces pol I and the nascent DNA becomes the leading strand of a newly formed replication fork.

The discovery of PriA has been instrumental in the study of recombination-dependent-replication (RDR) (see T4). Although it was initially thought not to be involved in *E.coli* chromosomal replication, because it is not necessary for primosome assembly at *oriC*, recent research supports the view, that PriA has a fundamental role in chromosomal DNA replication. Isolation of *priA* null mutant bacteria reveals a phenotype that is constitutively induced for the SOS DNA damage response. These bacteria are also sensitive to UV damage and double-strand breaks. In addition, they are defective in homologous recombination. Cell survival of *priA* mutants is also less than 50% of the wild-type. These phenotypes can now be explained through a recently realized activity of PriA, its ability to bind to D-loops. As mentioned above, D-loops are recombination intermediates generated from the invasion of a homologous strand of DNA, and, in addition to being involved in T4 DNA replication, are involved in double-strand break repair and homologous recombination. The first evidence for a connection between recombination and replication was provided by the analysis of the priA null bacteria which are unable to carry out these basic cellular functions.

III. Yeast DNA Replication

As indicated in the preceding sections, the extensive study of *E. coli* and bacteriophage DNA replication has yielded an enormous wealth of information regarding the mechanisms involved in DNA synthesis. However, *E. coli* being a prokaryotic organism has a cellular architecture that is fundamentally different from even the simplest eukaryotic organism. Eukaryotic organisms contain a nucleus, whereas *E. coli* do not. The possession of a nucleus permits the physical separation of the processes of translation and transcription from that of DNA replication. Also, the DNA genome of most eukaryotic organisms consists of several distinct chromosomes; these chromosomes most often contain several different origins of replication. Therefore, only studying *E.coli* and bacteriophage DNA replication will not tell us how eukaryotic cells regulate and organize their DNA replication process. With these concerns in mind, researchers needed to turn their attention to the specifics of eukaryotic DNA replication. Much of the study of eukaryotic DNA replication has been done on very simple eukaryotic microorganisms, such as yeast. Yeast cells rapidly divide and contain only 3.5 times more DNA than *E. coli* cells. Yeast genetic systems have also been well defined and serve as an experimentally powerful arm in the study of yeast DNA replication. It has been found that many of the types of enzymes and proteins, as well as the molecular mechanisms involved in the replication of the *E. coli* DNA genome, are also used in yeast DNA synthesis.

GLOSSARY

Circle-to circle replication: Replication of circular DNA initiated at a specific site or origin and proceeding bi-directionally yielding an exact copy of the original circular template.

Concatemers: A series of connected and tandomly repeated DNA sequences or genomes. Concatemeric circles are interlocked circular DNA sequences similar to the rings of a chain.

D-loop dependent DNA replication: DNA replication initiated at D-loop structures formed through the displacement of a strand of DNA by an invading complementary strand of DNA establishing a 3'OH terminus for DNA synthesis and a site for the recruitment of replication forks.

DNA Polymerase III Holoenzyme(DNA pol III HE): The DNA polymerase responsible for synthesis of a new daughter strand of DNA onto the template strand during the elongation phase of *E. coli* chromosomal DNA replication. Composed of two cores (proteins αεθ) connected by the γ complex (proteins τγδδ'χψ) with the addition of the processivity factor protein β.

Okazaki Fagments: Short stretches of 1000-2000 nucleotides synthesized during discontinuous DNA replication of the lagging DNA strand.

***OriC*:** The origin of chromosomal DNA replication in *E. coli*. Composed of minimal 245 bp region containing three AT-rich domains and four dnaA boxes.

Primosome: A group of protein responsible for priming the DNA during DNA replication. The primosome is responsible for lying down the initial RNA primers used as a 3'OH terminus for the elongation phase of *E. coli* chromosomal DNA replication and also for synthesis of the RNA primers that occur at the beginning of each Okazaki fragment.

R-loop dependent DNA replication: DNA replication initiated by the synthesis of an RNA primer via the action of an RNA polymerase (RNAP) through an AT-rich region of DNA, which then allows the binding of factors involved in the initiation events of DNA replication and serves as a 3'OH terminus for elongation to proceed.

RNA polymerase(RNAP): An enzyme responsible for the synthesis of RNA onto a DNA template. RNA polymerases can the products of the cells (cRNAP) or can be encode in a viral genome (vRNAP).

REFERENCES

1) Marians, K.J. (1992) Prokaryotic DNA replication. Annual Reviews of Biochemistry 61:673-719.

2) Bramhill, D., Kornberg, A. (1988) A model for initiation at origins of DNA replication. Cell 54:915-918.

3) Ogawa, T., Hirose, S., Okazaki, T., and Okazaki, R. (1977) Mechanism of DNA Chain Growth XVI. Analyses of RNA-Linked DNA pieces in *Escherichia coli* with polynucleotide kinase. J. Molec. Biol. 112:121-140.

4) Olson, M.W., Dallmann, H.G., McHenry, C.S. (1995) DnaX complex of *Escherichia coli* DNA polymerase III holoenzyme, the X.Ψ complex functions by increasing the affinity of τ and γ for δ.δ' to a physiologically relevant range. Journal of Biological Chemistry 270:29570-29577.

5) Leu, F.P., Hingorami, M.M., Turner, J., O'Donnell, M. (2000) The δ subunit of DNA polymerase III holoenzyme serves as a sliding clamp unloader in *Escherichia coli.* Journal of Biological Chemistry 275:34609-34618.

6) Kong, X.P., Onrust, R., O'Donnell, M., Kuriyan, J. (1992) Three-Dimensional Structure of the b Subunit of E. coli DNA Polymerase III Holoenzyme: A Sliding Clamp. Cell 69:425-437.

7) Kim, D.R., McHenry, C.S. (1996) Identification of the β-binding domain of the α subunit of *Escherichia coli* polymerase III holoenzyme. Journal of Biological Chemistry 271:20699-20704.

8) Bussiere, D.E.,Bastia, D. (1999) Termination of DNA replication of bacterial and plasmid chromosomes. Mol. Microbiol. 6:1611-8.

9) Krenzer, K.N. (2000) Recombination-dependent DNA replication in phage T4. Trends in Biochemical Science 25:165-173.

10) Marians, K.J. (2000) PriA: at the crossroads of DNA replication and recombination. Progress in Nucleic Acids Research 63:39-65.

Glossary

Circle-to circle replication: Replication of circular DNA initiated at a specific site or origin and proceeding bi-directionally yielding an exact copy of the original circular template.

Concatemers: A series of connected and tandomly repeated DNA sequences or genomes. Concatemeric circles are interlocked circular DNA sequences similar to the rings of a chain.

D-loop dependent DNA replication: DNA replication initiated at D-loop structures formed through the displacement of a strand of DNA by an invading complementary strand of DNA establishing a 3'OH terminus for DNA synthesis and a site for the recruitment of replication forks.

DNA Polymerase III Holoenzyme(DNA pol III HE): The DNA polymerase responsible for synthesis of a new daughter strand of DNA onto the template strand during the elongation phase of *E. coli* chromosomal DNA replication. Composed of two cores (proteins $\alpha\varepsilon\theta$) connected by the γ complex (proteins $\tau\gamma\delta\delta'\chi\psi$) with the addition of the processivity factor protein β.

Okazaki Fagments: Short stretches of 1000-2000 nucleotides synthesized during discontinuous DNA replication of the lagging DNA strand.

***OriC*:** The origin of chromosomal DNA replication in *E. coli*. Composed of minimal 245 bp region containing three AT-rich domains and four dnaA boxes.

Primosome: A group of protein responsible for priming the DNA during DNA replication. The primosome is responsible for lying down the initial RNA primers used as

a 3'OH terminus for the elongation phase of *E. coli* chromosomal DNA replication and also for synthesis of the RNA primers that occur at the beginning of each Okazaki fragment.

R-loop dependent DNA replication: DNA replication initiated by the synthesis of an RNA primer via the action of an RNA polymerase (RNAP) through an AT-rich region of DNA, which then allows the binding of factors involved in the initiation events of DNA replication and serves as a 3'OH terminus for elongation to proceed.

RNA polymerase(RNAP): An enzyme responsible for the synthesis of RNA from a DNA template. RNA polymerases can the products of the cells (cRNAP) or can be encode in a viral genome (vRNAP).

Figure 1. The *E.coli* chromosomal DNA replication origin, *oriC*. The four boxes to the right shaded gray represent the 9 base pair (bp) recognition sites for dnaA proteins, and the three boxes to the left represent the 13 bp AT-rich domains melted during the initiation of *E.coli* DNA replication.

Figure 2. Initiation of chromosomal DNA replication in *E.coli*. DnaA binds to the 9bp sequences of *oriC* in the presence of the protein HU and ATP forming a complex of 20-40 dnaA molecules around which the DNA of the origin is wrapped and is referred to as the initial complex. In the presence of 5mM ATP and at 38°C, the initial complex is converted to an open complex via the ATP-dependent melting of the AT-rich domain by dnaA. The dnaB-dnaC complex is then directed to the melted region of DNA through the interaction of dnaC with dnaA. DnaB is now able to recognize the single-stranded DNA further unwinding the DNA leading to the formation of the prepriming complex.

Figure 3. The *E.coli* replication fork. After initiation has occurred, two replication forks form at the points where single-stranded DNA meets double stranded DNA. (A) The diagram of the proteins present at a replication fork in *E.coli*. The progression of the replication fork is directed through the helicase activity of dnaB followed by the synthesis of RNA primers by dnaG on both the leading (bottom strand) and lagging strand (top strand). The DNA polymerase III holoenzyme (DNA pol III HE) then synthesizes new strands of DNA beginning with the RNA primers. The areas containing single-stranded DNA are stabilized through the binding of single-stranded DNA binding proteins (SSB). A more accurate representation of a replication fork is shown in (B).

Here, the DNA is unwound by dnaB at the leading end of the replication fork, and a nascent strand of DNA is synthesized on the leading strand (top strand) by a DNA pol III HE. Tethered to the DNA by a β subunit dimer, the DNA pol III HE moves along the leading strand of DNA. Synthesis of the leading strand (top strand) is coordinated with synthesis of the lagging strand (bottom strand) through the dimerization of two DNA pol III HE complexes by the γ subunit. The DNA of the lagging strand is looped around one DNA pol III complexes, and following primer synthesis and loading of a β dimer at the primer terminus by the γ complex, the DNA pol III HE synthesizes multiple short 1000-2000 bp Okazaki fragments onto the lagging strand. These Okazaki fragments will become ligated to one another forming a long contiguous strand of DNA. The direction of fork movement and leading and lagging strand movement are denoted with arrows.

Figure 4. The Chromosomal DNA replication terminus of *E.coli.* The replication terminus is located 180° around the *E.coli* chromosome from *oriC* and contains two sets of three directly repeated sequences called *ter* sites. At these *ter* sites, the Tus protein binds to the DNA and abrogates replication fork movement. However, the movement of one replication fork can only be stopped by a Tus protein binding to a single *ter* site that is in the proper orientation. This unidirectional or polar blockade of replication by Tus is represented by the sequences to the left and the leftward moving replication fork (top arrow), and the sequences to the right and the rightward moving replication fork (bottom arrow). When a replication fork encounters these sequences bound with Tus, an interaction between dnaB and Tus occurs, most likely resulting in inhibition of its helicase activity and cessation of replication fork movement.

Figure 5. The DNA replication of phage λ . Upon infection, the λ DNA is injected into the bacterium as a linear molecule. The linear λ chromosome then cyclizes through the hybridization of cohesive termini (*cos* sites) at the ends of the molecule (A). Circle-to-circle DNA replication initiates at a single replication origin *oriλ* (B) by transcription of an RNA primer through the origin from one of two promoters (i.e. $P_{R1}$and P_{R2}). Replication forks are formed, and circle-to-circle replication proceeds producing two daughter molecules (C). After a few rounds of circle-to-circle replication, λ switches to an alternative mode of DNA replication called rolling circle replication (E-G). Rolling circle replication initiates by the nicking of one strand of the circular DNA resulting in a free 3'-OH terminus that serves as a primer for DNA synthesis (E). Polymerization of the new strand, rolling round and round the circular molecule, causes displacement of the old strand forming long tandemly repeated λ chromosomes called concatemers (F). Single λ chromosomes are generated from the site-specific cleavage (lightning bolt) of the λ concatemer at the *cos* sites (G).

Figure 6. T7 DNA replication. DNA replication of the T7 phage initiates through the synthesis of an RNA transcript from one of two T7 promoters located to the left of the T7 origin of replication (*oriT7*), displacing the complementary DNA strand forming an R-loop (A). The 3'-OH terminus of the RNA transcript then serves as a primer terminus for leading strand synthesis, and two replication forks are assembled at the R-loop (B). DNA

replication then proceeds bi-directionally copying the T7 genome. The T7 genome, unlike λ, never cyclizes during its infection cycle, and. because the DNA polymerases require a 3'-OH terminus to elongate from, single-stranded DNA overhangs are generated after the removal of RNA primers from the 5' ends (C). To circumvent the loss of the DNA overhang at the ends of the T7 genome, the T7 phage forms long DNA concatemers by the hybridization of the terminal repeats (TR) (D). Full-length T7 genomes are then generated through cleavage of the concatemers at specific sequences with-in the TR called *pac* sites (E).

Figure 7. The mechanisms of DNA replication initiation from D-loops during T4 DNA replication. (A) Unidirectional DNA replication results from the cleavage of the hybridized (♦) strand of DNA followed by the formation of a single replication fork. (B) Bi-directional DNA replication occurs through the cleavage of the displaced (•) strand of DNA followed by formation of two replication forks. (C) Tri-directional DNA replication is the result of the formation of three replication forks on an uncleaved template. Alternatively, tri-directional DNA replication could also occur if two of the DNA strands were cleaved (not shown).

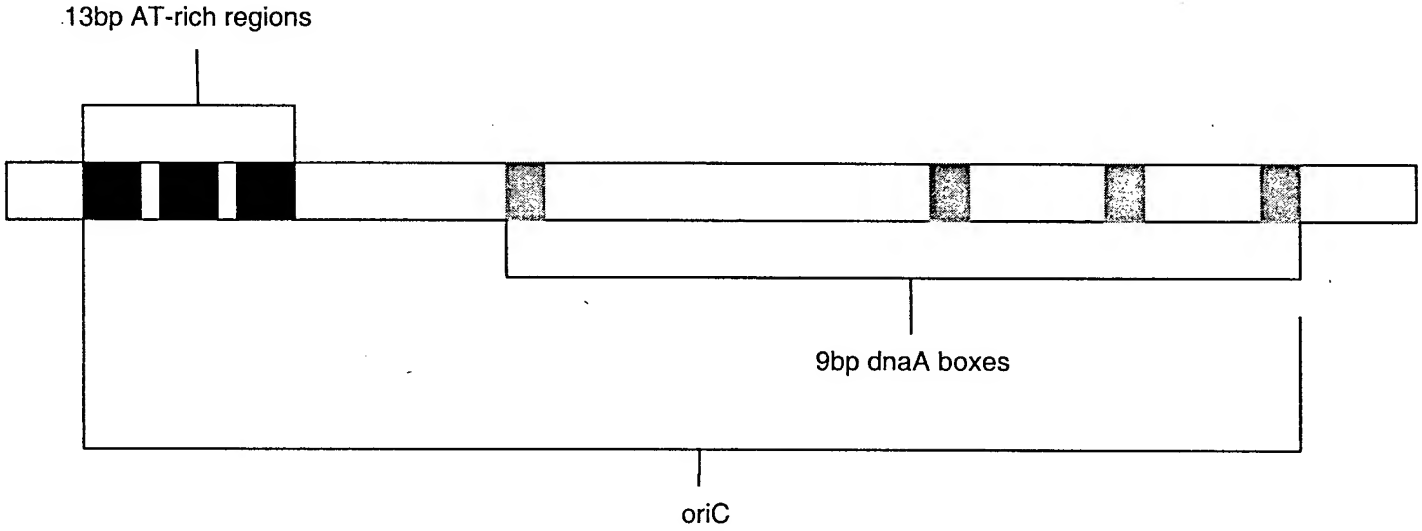


Figure 1

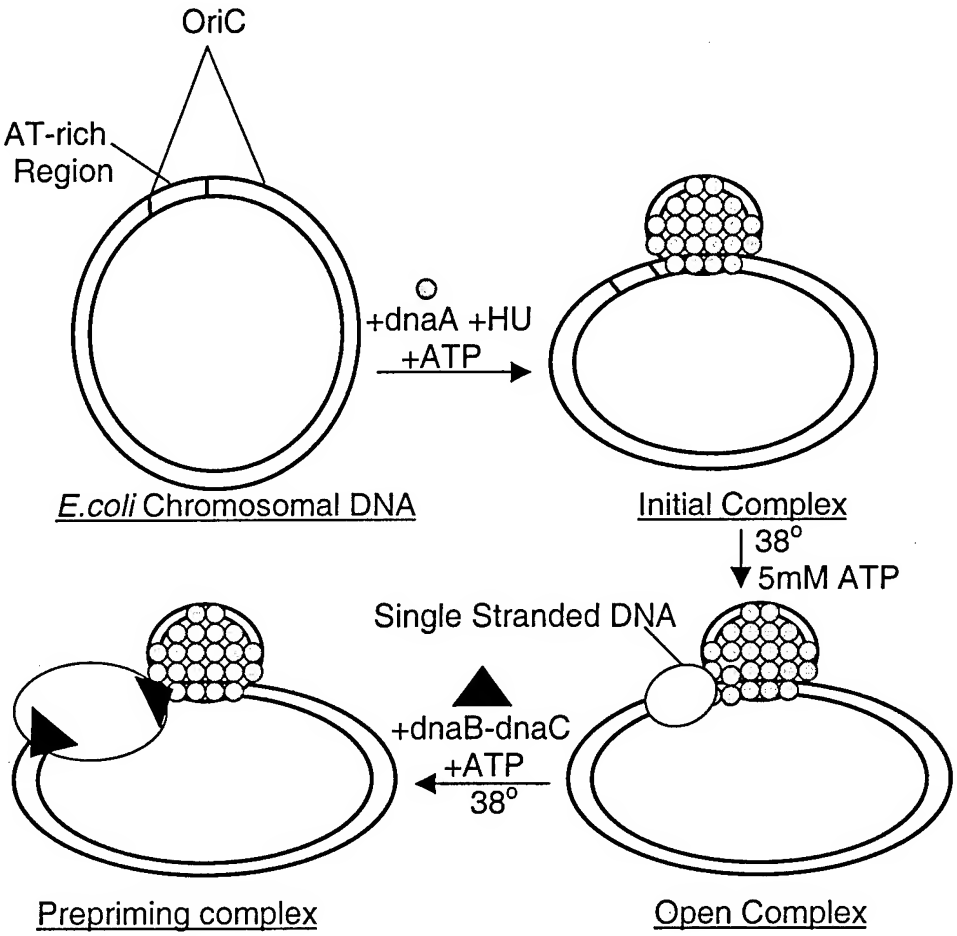
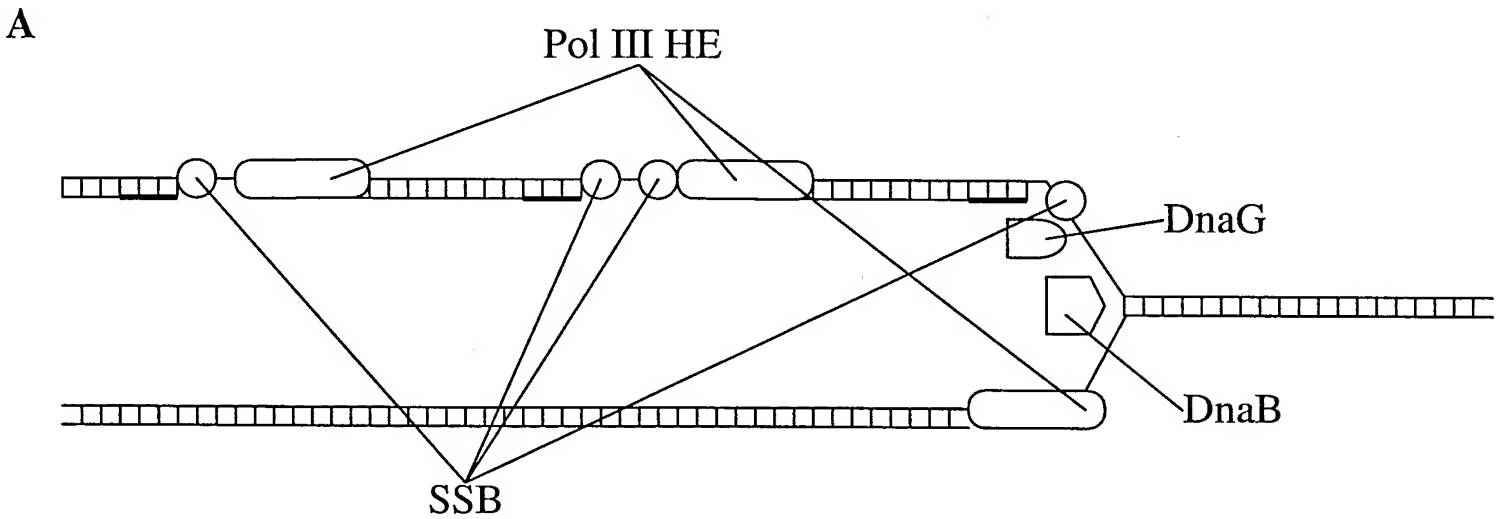

OriC
AT-rich
Region
+dnaA +HU
+ATP
E.coli Chromosomal DNA
Initial Complex
38°
5mM ATP
Single Stranded DNA
+dnaB-dnaC
+ATP
38°
Prepriming complex
Open Complex

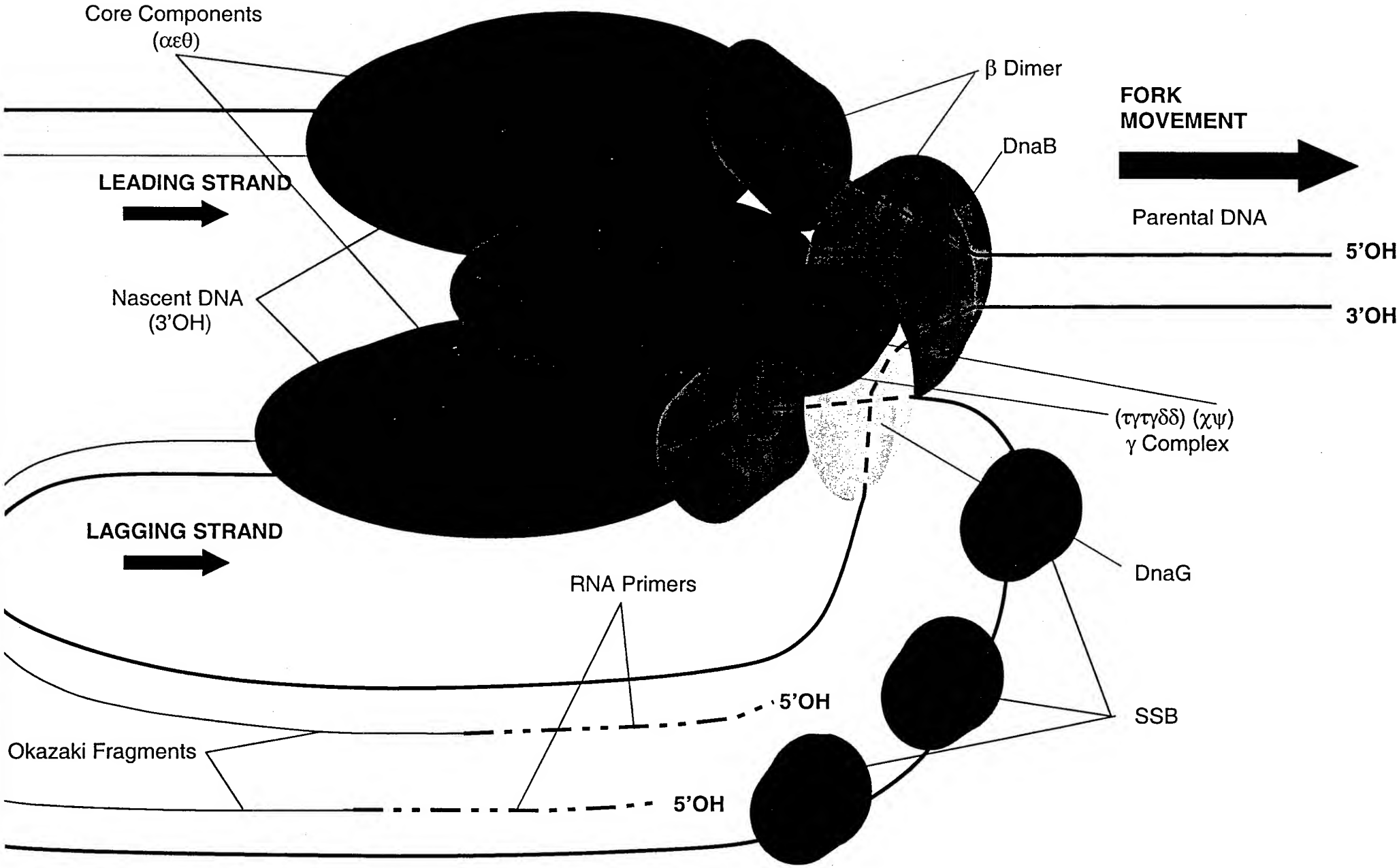
Figure 2

Figure 3A



B

Figure 3B


Core Components
(αεθ)
β Dimer
FORK
MOVEMENT
DnaB
LEADING STRAND
Parental DNA
5'OH
Nascent DNA
(3'OH)
3'OH
(τγτγδδ) (χψ)
γ Complex
LAGGING STRAND
DnaG
RNA Primers
5'OH
SSB
Okazaki Fragments
5'OH

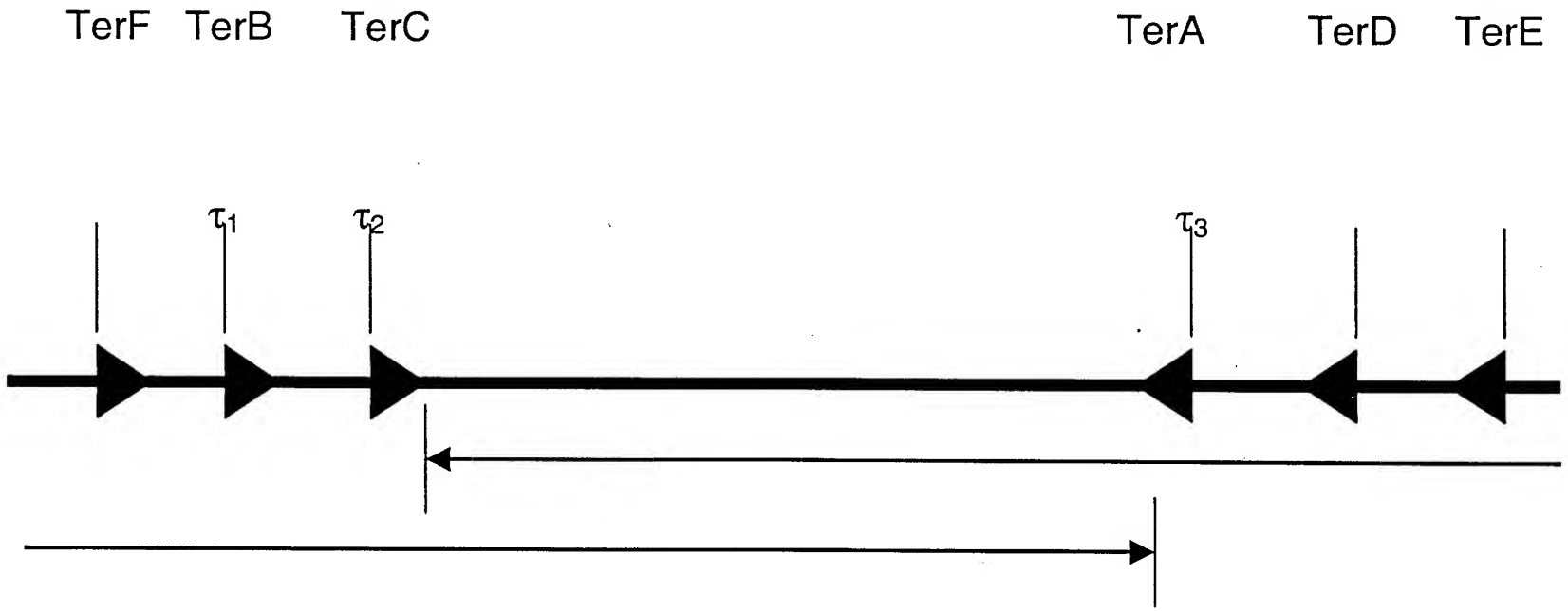


Figure 4

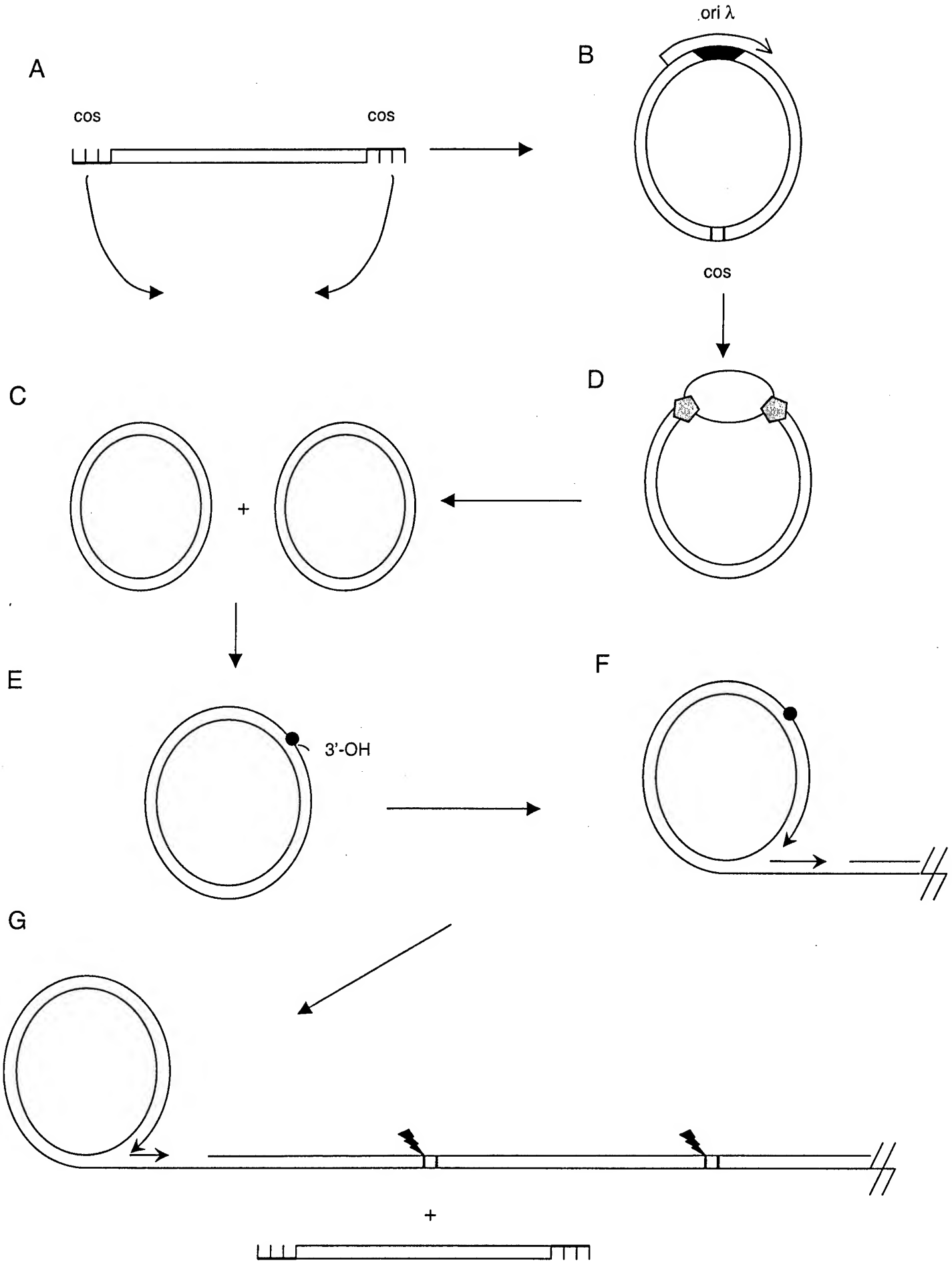


Figure 5

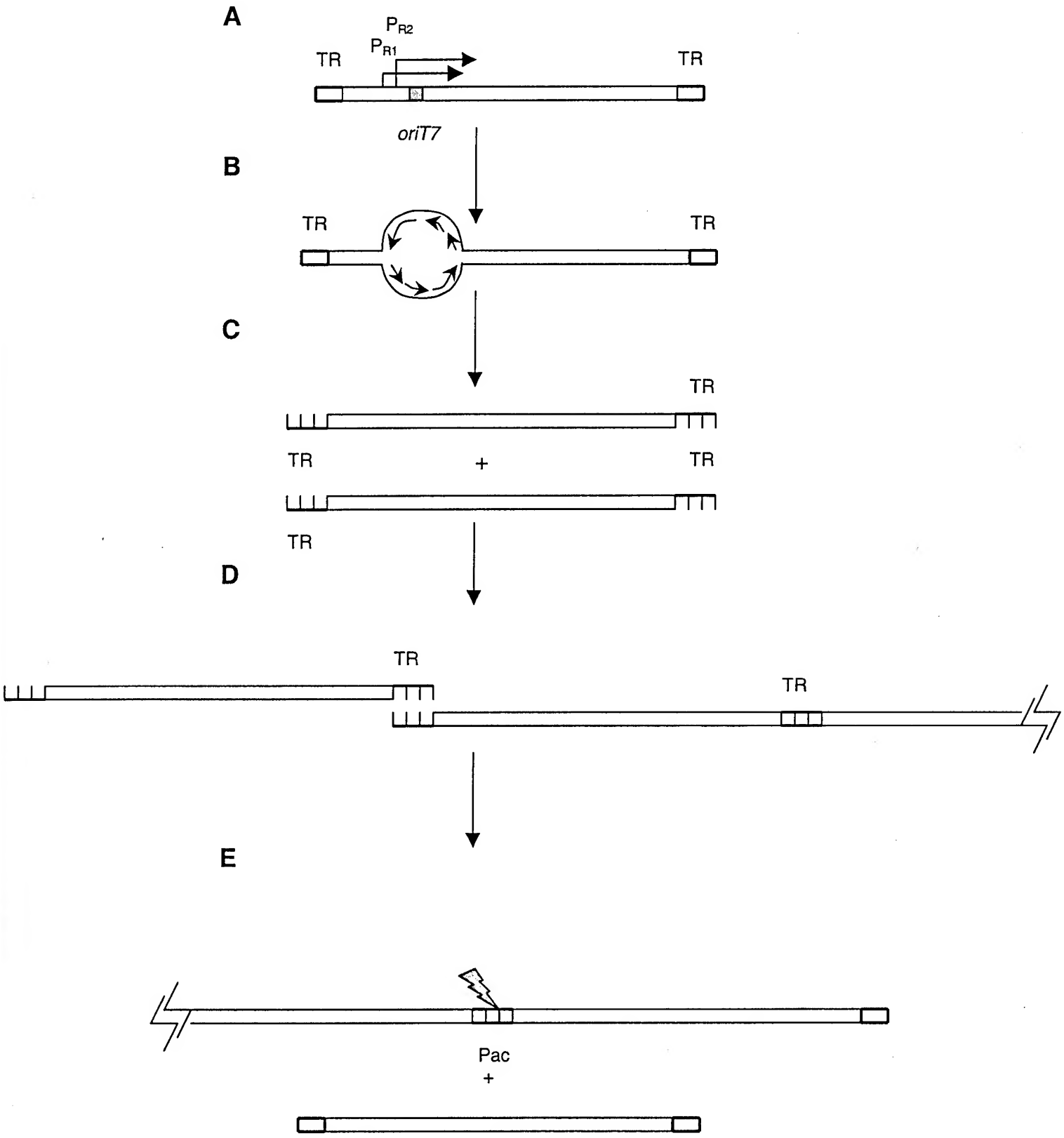


Figure 6

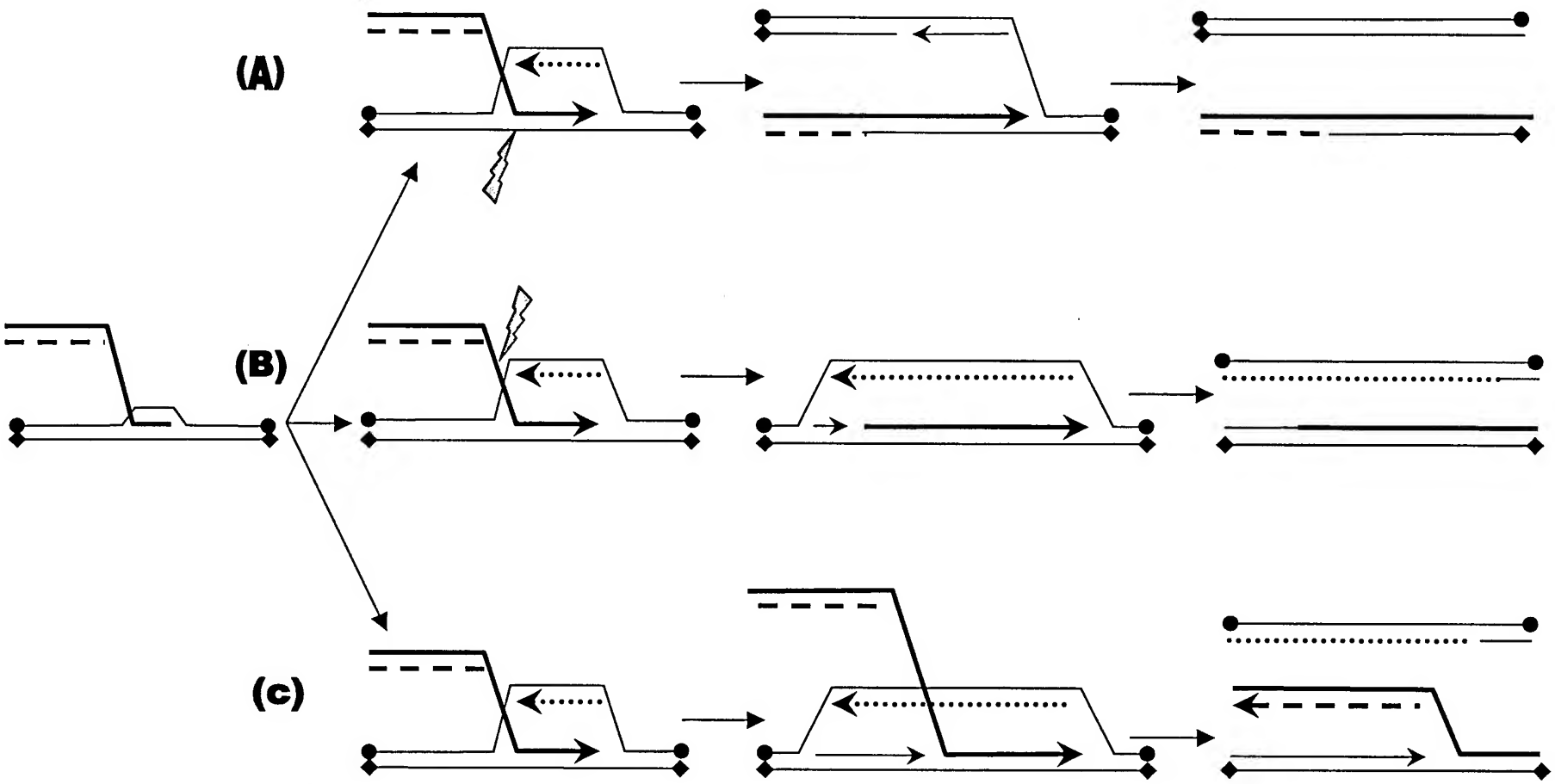


Figure 7

Title: **EUKARYOTIC DNA REPLICATION**

Authors: Linda H. Malkas, Derek J. Hoelz and Robert J. Hickey

Affiliation: University of Maryland School of Medicine
Department of Pharmacology and the Greenebaum Cancer Center
Health Sciences Facility I
685 West Baltimore Street
Baltimore, MD 21201

Phone: 410-706-2313

Fax: 410-706-0032

Email: lmalkas@som.umaryland.edu

I. MAMMALIAN CELL DNA REPLICATION

The process of DNA replication in mammalian cells is highly complex and has several unique features that distinguish it from simpler prokaryotic systems. First, chromosomal DNA replication is compartmentalized within the mammalian cell nucleus and is partitioned from the cytoplasm which is the site of synthesis of proteins and other metabolites that function in DNA synthesis, as well as the site for the mediation of extracellular stimuli that may trigger the initiation of DNA replication. Second, the mammalian chromosome is a complex nucleoprotein structure composed of both DNA and protein. These chromosomal-associated proteins must be duplicated along with the DNA in order to maintain proper chromosomal organization that in turn influences gene expression. Third, the mammalian chromosome contains multiple replication origins per DNA molecule that promote the initiation of DNA synthesis in a precise and temporally regulated manner. The precise firing of these replication origins leads to a spatially regulated chromosomal DNA synthesis occurring in defined replication units or replicons. Therefore, in that mammalian cells contain multiple chromosomes as well as many replication origins on a single chromosome, the act of chromosomal DNA replication must be a highly coordinated process.

To successfully mediate the complicated coordination of mammalian DNA replication it is reasonable that the process does not occur by random collisions between soluble enzymes and substrates but rather that it is orchestrated by an organized multiprotein complex that could be thought of, in effect, as a molecular machine. Molecular machines have been previously identified for the processes of protein synthesis and RNA transcription. However, it has only been in recent years that strong evidence for organized interactions between mammalian DNA replication proteins has been described.

In this article the current knowledge regarding the proteins mediating mammalian DNA replication is described. There is also a discussion of the role the mammalian cell nuclear architecture plays in DNA replication, as well as the current findings regarding mammalian DNA sequences that potentially serve as origins of DNA replication. The

evidence for the existence of an organized DNA replication machine in mammalian cells is also presented.

I A. Proteins That Drive the Mammalian DNA Replication Fork

The study of mammalian DNA replication lagged behind that of prokaryotes for many years. This was due to the lack of a reliable and efficient mammalian cell-based *in vitro* DNA replication system. In 1984 the first mammalian-based DNA replication system that successfully initiated DNA synthesis *in vitro* was developed by Li and Kelly. The replication system exploited some of the unique aspects of a mammalian virus, simian virus 40 (SV40), to drive the system.

SV40 is a small DNA tumor virus and a member of the papovavirus family. The virus replicates its genome in the nucleus of its permissive host cell and requires only two viral components. All other replication factors necessary for SV40 DNA synthesis are provided by the host. Because the viral DNA replication is almost completely dependent on the host cell DNA replication machinery, *in vitro* papovavirus DNA replication is not only useful for the study of viral DNA synthesis, it has also proven to be invaluable for the investigation of mammalian DNA replication as well. To date, it is still the best model available for the study of mammalian DNA synthesis *in vitro*. One of the required viral components for the SV40-based *in vitro* DNA synthesis system is the viral replication origin. The other component is the only virally encoded protein required for DNA replication, the large T-antigen. The large T-antigen recognizes and binds to the viral replication origin. The large T-antigen also has a helicase activity that melts the DNA in the replication origin that then presumably allows access of replication proteins to this DNA.

A number of reviews have been written over the years detailing the development of the mammalian DNA replication field following the advent of the SV40-based *in vitro* DNA synthesis system, and readers may wish to consult them. However, for the purposes of this article the employment of the SV40-based *in vitro* DNA replication system has led to the identification of several mammalian DNA replication proteins. The proteins that

were demonstrated to be required for *in vitro* DNA replication include: DNA polymerase α-primase, DNA polymerase δ, proliferating cell nuclear antigen (PCNA), replication protein A (RPA), replication factor C (RFC), DNA ligase I, topoisomerases I and II, and nucleases RNase H1 and FEN1/RTH1. The proposed action of these proteins at the mammalian DNA replication fork is shown in Figure 15 and described in more detail below. These proteins represent the minimal number of factors needed to drive the mammalian DNA replication fork.

I B. Proteins Mediating Leading Strand Synthesis

As seen previously for *E. coli* DNA synthesis, the antiparallel nature of duplex DNA together with the unidirectional mechanism of action of the known mammalian replicative DNA polymerases (i.e., as in *E. coli*, these enzymes only polymerize in the 5'→3' direction) dictates that one new daughter DNA strand be synthesized in a continuous fashion in the same direction as fork movement (designated the leading strand). The other daughter strand, therefore, is replicated opposite to the direction of fork movement and designated the lagging strand. As described for *E. coli*, the lagging strand is synthesized as short discontinuous DNA segments, defined as Okazaki fragments. Each new Okazaki fragment is initiated from a short RNA primer. As replication continues the RNA primer is excised, the oligonucleotide gap is filled in by a DNA polymerase and finally each Okazaki fragment is enzymatically joined (or ligated) to complete the synthesis of the lagging DNA strand.

Our current information on how the leading and lagging DNA strands are synthesized in mammalian cells has been derived almost entirely from the use of the SV40-based *in vitro* DNA replication system. The current model for leading strand DNA synthesis is as follows. The duplex DNA in an origin of replication is melted by the action of an as-yet-to-be-defined mammalian helicase to separate the DNA strands to create two replication forks. A single-stranded DNA binding protein (RPA) binds the melted DNA to maintain the opened duplex DNA structure, and promote further unwinding; thereby presumably allowing access of other replication proteins to the DNA. After the initial DNA unwinding, leading strand synthesis is commenced by the synthesis of a short RNA

primer. This primer is synthesized by the action of primase which is tightly associated with the enzyme DNA polymerase α. DNA polymerase α then begins to add deoxyribonucleotides onto the RNA primer. The replication protein RFC is then proposed to mediate a process termed "polymerase switching". RFC facilitates the dissociation of DNA polymerase α from the growing leading strand and promotes the loading of PCNA onto the DNA. ATP is expended in this process. Three molecules of PCNA are loaded onto the DNA in such a manner as to form a clamp around the DNA strand. DNA polymerase δ then associates with PCNA, making the polymerase highly processive and permitting the addition of thousands of deoxynucleotides to the growing leading strand in one binding event.

I C. Proteins Mediating the Synthesis of the Lagging DNA Strand

Current evidence derived from the SV40-based *in vitro* replication model system suggests that priming by DNA polymerase α and subsequent switching to DNA polymerase δ occur similarly on both the lagging and leading daughter strands during DNA synthesis. However, there must be frequent polymerase switching to accommodate the synthesis of Okazaki fragments on the lagging daughter DNA strand. Two nucleases are required to excise the many RNA primers that result during lagging strand synthesis. These nucleases are RNase H1 and FEN1/RTH1. The joining of adjacent Okazaki fragments to complete lagging strand synthesis is mediated by DNA ligase I.

I D. Roles for Topoisomerases I and II in DNA Replication

DNA topoisomerases I and II have been suggested to play a role in *in vitro* SV40 DNA replication. I*n vitro* SV40 DNA replication activity was restored by adding purified topoisomerases I and II to depleted cell extracts. In addition, it was observed that either the topoisomerase I or II enzyme could carry out the unwinding activity required for the progression of the replication fork. However, a unique role was also described for topoisomerase II in SV40 DNA synthesis *in vitro*. It was observed that topoisomerase II can also function in the segregation of newly replicated daughter molecules.

These *in vitro* SV40 DNA replication studies support the work of others, who, using hypotonic shock of mammalian cells, provided the first evidence to show that in intact cells topoisomerase II is required to separate daughter chromosomes. It was then demonstrated conclusively that topoisomerase II was required to separate daughter chromosomes in yeast. Additional studies using intact mammalian cells showed that topoisomerase II inhibitors effectively block SV40 chromosomal decatentation. These intact cell studies provide the strongest evidence that DNA topoisomerase II is required for mammalian cell DNA replication.

II. DNA REPLICATION AND MAMMALIAN CELL NUCLEAR ARCHITECTURE

Compelling evidence on the critically important role of nuclear architecture on mammalian DNA synthesis has been reported. Highlights on concepts regarding nuclear architecture as it pertains to the nuclear matrix and the organization of what are termed "DNA replication factories" are described in the following sections of this article.

II A. The Nuclear Matrix

Ultrastructure studies of the mammalian cell's nucleus have shown an association of both euchromatin and heterochromatin with an extensive nonchromatinous filamentous and granular network in the interior of the nucleus. Over 50 years ago it was found that a fraction of a cell's nuclear proteins readily resisted extraction even after treatment with buffers of high ionic strength. This nuclear subfraction was shown to be made of the nucleoprotein fibrillary network. The term "nuclear matrix" was applied to this subcellular structure in 1974 by Berezney and Coffey. After much work the nuclear matrix is now mostly considered a valid subcellular structure with several currently defined functions and most likely a host of duties yet to be identified.

It was suggested that the nuclear matrix may participate in the unwinding of the DNA double helix to allow access of the DNA synthetic machinery and to permit the easy separation of daughter DNA molecules from the parental template during replication. However, conclusive evidence showing a matrix binding region on the daughter DNA

molecule or specific attachment sites for the DNA on the nuclear matrix has not been demonstrated.

Other evidence strongly implicates the nuclear matrix in chromosomal DNA replication. Indirect immunofluorescence microscopy studies on human cells using a monoclonal antibody directed against DNA polymerase α revealed a granular pattern of fluorescence in the nuclei of the cells. The fine granules and meshwork of the fluorescence resembled the pattern that is characteristic of the nuclear matrix. In addition, fluorescent microscopy was used to map DNA replication sites in the mammalian interphase cell nucleus after incorporation of biotinylated dUTP. Discrete replication granules were observed to be distributed throughout the nuclear interior and along the periphery. *In situ* prepared nuclear matrix structures also incorporated biotinylated dUTP into replication granules that were indistinguishable from those detected within the intact nucleus. It was suggested that each replication granule may correspond to a replication cluster assembly in which numerous tandemly arranged replicons are coordinately synthesizing DNA.

In other work, the retention of a large megacomplex containing DNA polymerase α and other enzymes for DNA synthesis was reported in agarose-entrapped nuclei. The megacomplex could function in the replication of endogenous chromosomal DNA, and some evidence was obtained suggesting that it associates with the nuclear matrix. It was also reported by others that 100 to 150S megacomplexes containing DNA polymerase α-primase were associated with the nuclear matrix from regenerating rat liver. The megacomplexes were believed to be composed of clusters of 10S and 17S protein complexes that underwent dissociation when the isolated megacomplexes were left at 4°C for one hour or more. Although more extensive investigation is required, these studies strongly suggest that the mammalian DNA synthetic apparatus assembles in some fashion into megacomplexes. These megacomplexes then associate with the components of the nuclear architecture to form a supramolecular structure for DNA synthesis.

There is a tantalizing correlation between the sizes reported for the rat liver nuclear matrix associated DNA polymerase α-primase containing complex (17S) and that for the

replication-competent murine cell DNA replication complex (synthesome) isolated in our own laboratory, which will be described in more detail below. The murine DNA synthesome is also 17S and fully supports *in vitro* papovavirus DNA replication. The relationship of these complexes to one another is currently unknown.

Additional evidence in support of a role for the nuclear matrix in mammalian cell DNA replication was provided by researchers who showed the position of a DNA replication origin relative to the nuclear matrix by the autoradiographic analysis of nuclear matrix halo structures. In synchronized BHK cells it was observed that labeled DNA at the beginning of S-phase remained matrix associated, and that it later would migrate into the DNA halo; suggesting that replication origins remain matrix bound after the initiation of DNA synthesis.

A variety of DNA and RNA metabolism proteins have been reported to be nuclear matrix associated. These have included DNA polymerases α and β, topoisomease II, terminal deoxyribonucleotidyl transferase, RNA polymerase II, DNA methyl transferase, and DNA primase. Although the function of each of these proteins is known, the how and why of their nuclear matrix association remains a mystery.

II B. DNA Replication Factories

DNA synthesis in higher eukaryotes appears to occur within replication foci. These replication sites contain many replicons, and within them DNA synthesis is mediated by a single megacomplex of proteins. The replication factories apparently assemble in a cell-cycle-dependent manner, appearing initially in the late G1 phase and are maintained throughout S phase. It was observed that characteristic patterns of synthesis could be defined, suggesting a specific S phase program for replication. These data suggest that in higher eukaryotes a temporally coordinated activation of different sets of replicons occurs throughout the S phase, and that activation of these sites is influenced by nuclear structure. Early in the S phase the factories measure approximately 100 nm across, and each factory was estimated to duplicate 10 replicons. The high level of organization of the replication factories can be envisioned to allow a great deal of regulatory control over

the act of DNA synthesis. The regulation of the replication factories remains undefined at this time. Also, the relationship between the replication factories and the mammalian cell DNA synthesome that will be described below is currently unknown.

II C. A Potential Regulatory Role for the Nuclear Matrix during the Initiation of DNA Synthesis

As stated above, the nuclear matrix is a highly organized structure around which DNA and RNA synthesis is coordinated, and DNA replication appears to be carried out in "replication factories" that appear to be associated with the nuclear matrix. Evidence consistent with this suggestion has been presented by a number of other investigators. Functional evidence supporting this concept has demonstrated that the earliest replicating DNA is associated with the nuclear matrix, and that only nuclear matrix packaged DNA was able to correctly initiate DNA synthesis from a specific replication origin sequence. Thus far, the only matrix attachment consensus sequences observed in higher eukaryotes are the sequence ATTA and the closely related ATTTA sequence. Inverted repeats are also found in mammalian replication origins; however, the presence of cruciform structure is not detectable at matrix attachment sites. The data provided by these laboratories suggest the importance of the nuclear matrix in organizing the DNA and in potentially regulating the initiation of DNA synthesis at a specific DNA sequence. However, evidence has also been presented indicating that the recognized matrix attachment consensus sequences are not essential elements in all higher eukaryotic replication origins isolated thus far. This observation does not imply that matrix attachment is not important to the proper regulation of DNA synthesis, because unrecognized matrix attachment regions might be present in the DNA adjacent to these isolated replication origins.

III. THE ISOLATION OF A FULLY FUNCTIONAL DNA REPLICATION MACHINE FROM MAMMALIAN CELLS

The first successful isolation, extensive purification and characterization of an intact mammalian cell multiprotein DNA replication complex that is both stable and fully functional *in vitro* has been described. The replication-competent multiprotein complex

has been isolated from a wide variety of mammalian cells grown in culture as well as from tissue. The multiprotein DNA replication complex has been purified approximately 3600-fold from cells using a series of centrifugation, polyethylene glycol precipitation, ion-exchange chromatography, density gradient sedimentation and native polyacrylamide gel electrophoretic steps. The sedimentation coefficient of the multiprotein complex from human cells was 18S, as determined by sucrose gradient analysis, while that isolated from murine cells was 17S. It has also been shown that the replication complex was a discrete protein species in 4% native polyacrylamide gels. The integrity of the multiprotein complex was maintained after its treatment with detergents, salt, RNase, DNase, chromatography on anion exchange resins, sedimentation in glycerol and sucrose density gradients, and electrophoresis through native polyacrylamide gels, indicating that the association of the replication proteins with one another was independent of non-specific interaction with other cellular macromolecular components. Most importantly, it was demonstrated that this complex of proteins was fully competent to replicate DNA *in vitro*. The human cell complex supported *in vitro* SV40 DNA synthesis, while the mouse cell complex replicated polyomavirus origin containing duplex DNA. This result indicated that all of the cellular activities required for large T-antigen dependent *in vitro* papovavirus DNA synthesis were present within the isolated multiprotein form of the DNA replication apparatus. This is particularly important in view of the fact that papovavirus is completely dependent on the host cell's DNA synthetic machinery for its own DNA replication. These results, therefore, indicate that the isolated multiprotein complex mediates not only papovavirus DNA synthesis, but it must function in mammalian cell DNA replication as well.

The DNA replication proteins identified to co-purify with the multiprotein complex, thus far, are: DNA polymerase α-primase, DNA polymerases δ and ε, DNA ligase I, RPA, RFC, PCNA, and topoisomerases I and II. A DNA helicase activity was also observed to co-purify with the replication-competent multiprotein complex. This replication complex associated DNA helicase activity may function to promote the melting of parental DNA strands to permit access of the elongation components of the replication complex to the DNA.

It has also recently been shown that poly (ADP-ribose) polymerase (PARP) co-purifies with both the human and murine cell derived replication complex. PARP has been demonstrated to modulate the structure and function of a number of proteins involved in DNA metabolism through their poly (ADP) ribosylation. It was found that approximately 15 of the 35 polypeptides composing the replication complex are poly (ADP) ribosylated. These data strongly suggested that poly (ADP) ribosylation of some the replication complex's components may modulate its DNA synthetic activity.

The mammalian cell replication complex was also examined for its ability to utilize nucleosome-assembled duplex DNA as a parental DNA replication template. The experiments indicated that nucleosome containing DNA was replicated efficiently by the replication complex, indicating that the complex was fully capable of synthesizing DNA that was assembled into a chromatin-like structure.

A model to represent the mammalian Multiprotein DNA Replication Complex (designated the MRC or DNA synthesome) has been described and was based on the observed fractionation, chromatographic and sedimentation profiles of the individual replication proteins found to co-purify with the complex, Figure 16. The proteins, DNA polymerases α, δ, and ε, DNA primase, RF-C, PARP, DNA helicase, DNA ligase I and topoisomerase II, are observed to exclusively co-purify with one another during the isolation of the DNA synthesome. This would suggest that these proteins are "tightly" associated with one another, and it was proposed that these proteins form the "core" of the DNA synthesome. It is interesting to note that these proteins also function primarily in the elongation phase of DNA synthesis. PCNA, which functions as an accessory factor for DNA polymerase δ, was observed to be more "loosely" associated with the DNA synthesome. This may reflect this protein's suggested diverse functions in both DNA replication and repair. PCNA was represented in the model as a component of the DNA synthesome, but not as a member of the core.

Topoisomerase I and RP-A, together with the helicase activity of the papovavirus large T-antigen, are involved in the initiation events of papovavirus DNA replication. A similar role for the topoisomerase I and RP-A proteins associated with the mammalian DNA synthesome has been proposed. It was observed that topoisomerase I and RP-A, like PCNA, do not "tightly" associate with the DNA synthesome protein members that compose the core. It was therefore proposed that topoisomerase I and RP-A constitute the DNA synthesome's "initiation" component. Together, the core and initiation components constitute the synthesome.

Recently, additional evidence was obtained indicating the direct protein-protein interaction of DNA polymerases α and δ, DNA primase, RF-C and PCNA with each other in the DNA synthesome. These results were supported by previous work of other laboratories. These findings have led to a new model for the organization of several of the proteins within the core component of the DNA synthesome, Figure 17. Overall, the data suggested that the mammalian cell DNA synthesome represents an asymmetric multiprotein complex for DNA replication. According to the current model for DNA replication fork movement, described above, polymerase α-primase synthesizes RNA primers required for the initiation of leading DNA strand and Okazaki fragment synthesis, whereas, DNA polymerase δ conducts both leading and lagging strand DNA synthesis during the elongation phase of DNA replication. It has been postulated that RF-C facilitates coordinated leading and lagging strand synthesis by serving as a molecular hinge between DNA polymerases α and δ. Studies performed with the DNA synthesome support this hypothesis as they provided direct evidence for a close association between DNA polymerases α, δ and RF-C within the complex, Figure 17. Additionally, PCNA may participate in the coordination of leading and lagging strand synthesis by the DNA synthesome, as it was found to tightly associate with DNA polymerases α, δ and RF-C. Such a role for PCNA in DNA replication is supported by studies demonstrating that only early replicative lagging strand products are synthesized *in vitro* in the absence of the protein. Furthermore, as DNA polymerase ε is a core component of the mammalian cell DNA synthesome and closely associates with PCNA in the complex, it suggests that this protein also may play a role in DNA replication.

Further work must be performed in order to determine the exact function of polymerase ε within the DNA synthesome. However, it has been proposed that the protein may mediate the conversion of DNA primers into Okazaki fragments or act as a molecular sensor of DNA damage in eukaryotic cells. The protein-protein interactions of synthesome-associated DNA ligase I and topoisomerases I and II have yet to be established.

IV. MAMMALIAN DNA REPLICATION ORIGINS

Mammalian origins of DNA replication are highly complex structures composed of nucleotide sequences. They encode specific elements required to destabilize the helical structure of the DNA, bind specific protein factors that participate in the initiation phase of the DNA synthetic process, and specify the position within a chromosome from which DNA replication will initiate.

IV A. General characteristics of nucleotide sequences serving as replication origins

The nucleotide sequence encoding several bacterial, bacteriophage and eukaryotic viral origins of DNA replication have been clearly defined and their structures essentially elucidated. Considerable progress has also been made in identifying and characterizing the structure of replication origins in the budding yeast *S. cerevisiae*. In contrast to these origins of replication, mammalian cell replication origins appear to be far more complex structures; consisting of diverse sets of core origin nucleotide sequences, (i.e., replication initiation domains), DNA unwinding elements (DUE), matrix attachment regions, and transcription factor binding sites. Only a few of the replication initiator sites, at which mammalian chromosomal replication originates, have been identified, and the complete characterization of these and other potential initiator sites is still underway. The limited supply of information describing the characteristics of mammalian cell origins of replication contrasts sharply with that available on bacteria, bacteriophage and eukaryotic viral replication origins. This is due in part to the apparent dependence of mammalian origins on origin structure and topology rather than on highly conserved, specific nucleotide sequences. This feature of mammalian cell replication origins has made their isolation and subsequent characterization an extremely difficult task.

IV B. DNA replication initiates within specific chromosomal domains

In mammalian cells, DNA synthesis appears to initiate within small chromosomal domains rather than at specific conserved nucleotide sequences. There appear to be multiple preferred (or primary) sites of initiation within each chromosome, as well as a variety of secondary initiation sites. These secondary sites appear to be utilized to a far lesser degree than the primary initiation sites, and it is possible that these secondary sites direct the initiation of DNA synthesis when the adjacent primary site becomes inoperable. The degree to which one initiation site is chosen over another is at present thought to be due to a combination of factors that influence origin recognition and utilization. These factors appear to involve the organization of the structural elements recognized as components of a replication origin. In addition, the organization of the elements within a chromosome is believed to participate in the structure and topology of the mammalian replication origin; suggesting that these elements have the ability to participate in the regulatory processes which can control the initiation of DNA synthesis.

IV C. Current models describing mammalian replication origins

Several models have been proposed in an attempt to reconcile the apparent discrepancy between the findings of some investigators that mammalian DNA synthesis initiates at specific DNA sequences and other results indicating that DNA replication can initiate at seemingly random points within the genome. The first model, termed the "Jesuit" model, postulates that there are a larger number of replication origins within the mammalian genome than are necessary to ensure its complete replication. Origin utilization is viewed as a stochastic process where initiator protein binding, DNA unwinding, and synthesis of RNA primers all contribute to a final probability of utilizing a specific replication origin. A second model, called the "strand separation" model, postulates the existence of a broad unstable initiation zone. Microbubbles form within this zone, and these bubbles result in the initiation of DNA synthesis from multiple sites within this zone. Initiation is accompanied by the unwinding of a large region of the chromosome. A third model, designated the "unidirectional bubble" model, proposes the presence of primary initiation sites, which are adjacent to weak secondary sites. Utilization of these secondary sites to

initiate DNA synthesis is thought to result from stresses created in the duplex DNA. Finally, a fourth model, described as the "reformation" model predicts that polymerase α-primase primes multiple initiation sites within a broad replication origin zone. The priming of the initiation site is followed by the switching of DNA polymerase δ for polymerase α within an organized replication factory, engaged with the DNA at the replication origin. The switching of DNA polymerases is proposed to be a necessary step; enabling the primed initiation sites to be elongated rapidly, and to permit potential mismatches in the newly replicated DNA to be corrected efficiently.

IV D. DNA unwinding elements

DNA unwinding elements (i.e., DUE's) appear to be located within or adjacent to the mammalian replicator, or "core-type" origin, sequences. These DUE's are AT rich, and have been identified in a variety of eukaryotic cellular and viral origins. The DUE is composed of a non-conserved nucleotide sequence, which exhibits reduced helical stability, relative to that of the surrounding DNA. Several cloned human DNA fragments containing DUE's exhibit a high degree of DNA synthetic activity. Sequence analysis of three of these cloned human DUE's indicates the presence of long stretches of dinucleotide repeats composed principally of the dinucleotides AT, and occasionally of the dinucleotides GT. DNA melting is facilitated by negative supercoiling within the DUE, and the association of single-stranded binding protein (e.g., RP-A) with the individual strands of the DUE. The association of RP-A with the DUE, following melting of the element, may be related to the observed loading of RP-A onto newly replicated DNA during M phase. Base stacking interactions within the DUE contribute significantly to the characteristics of the element; indicating that the characteristic helical instability of the DUE is not solely a function of the AT content of the origin. DUE sites are generally observed to lie adjacent to replication initiator protein binding sites. It has been suggested that this observation indicates that a replication fork can process in only one direction immediately following initiation. However, an alternate interpretation would suggest that unwinding of the DUE following initiator protein binding could create a microbubble structure at the DUE facilitating access of replication complex

components to the DNA forks formed at either end of the microbubble. This would enable replication to proceed bi-directionally following unwinding of the DUE.

The helical instability of the DNA unwinding element may also have a role in creating alternate secondary structure such as Z-DNA, triple-stranded DNA, bent DNA or DNA cruciforms. Consistent with this suggestion is the observation that AT-rich regions are commonly associated with nuclear matrix attachment regions, chromosome scaffold attachment sites, topoisomerase II recognition sites, and replication initiator protein binding sites. In addition, the identification of specific proteins that bind AT-rich sequence tracts within replication origins has been reported, and there is at least one report that these AT tract binding proteins may be associated with a distinct DNA helicase.

IV E. Protein interactions with specific mammalian replication origin nucleotide sequences

A variety of proteins that bind to specific DNA sequences associated with mammalian replication origins have been reported. These proteins range in size from 28 to 150 kDa, and recognize nucleotide sequences that are generally AT rich. One notable exception is the PUR protein which binds to a GC-rich consensus element found near a number of replication origins and promoter elements. Several years ago a six polypeptide protein complex termed the origin recognition complex (i.e., ORC) was identified through a series of studies designed to identify proteins which specifically recognize a yeast origin of DNA replication. Subsequent studies have demonstrated the importance of these polypeptides in maintaining the proliferative ability of the yeast cell. Recently, the human counterparts to some of the yeast ORC subunits has been cloned. The identification of these human homologs to components of the yeast ORC suggests that human cells most likely contain an analogous complex that is likely to function in mediating events associated with transcription and potentially DNA replication. It also suggests that there must be DNA sequence elements that are related to one another in both the yeast and human genomes. These sequence elements are presumably recognized by the corresponding ORC complex polypeptides. Presumably, the human homolog to

the yeast ORC can also be expected to mediate transcription silencing and origin recognition.

IV F. Auxiliary factors facilitating the initiation of DNA synthesis

Transcription factor binding sites are now recognized as cis-acting auxiliary components which facilitate the initiation of DNA synthesis from mammalian replication origins. Four models have been proposed to explain how transcription factor binding facilitates the initiation of DNA synthesis from a replication origin. The first model suggests that origin-associated transcription factors promote the transcription of RNA molecules which act as primers. These primers are used during the initiation phase of the DNA synthesis process. The second model proposes that transcription factors act by: 1) preventing chromatin structure mediated repression of origin function. This model suggests that chromatin structure is maintained by displacement of histones from the DNA; or 2) counteracting the condensation of DNA into heterochromatin. The third model suggests that transcription factors can direct the binding of an initiator protein to a replication origin core sequence, and subsequently facilitate the association of the DNA synthetic apparatus with the initiator protein. The fourth model proposes that transcription factors activate the initiator protein(s), rather than recruit this protein(s) to the origin. Activation of the initiator protein then facilitates unwinding of the DNA, which in-turn promotes primer synthesis.

The evidence presented for the loading of the eukaryotic single-stranded DNA binding protein, RP-A, onto eukaryotic viral origins by transcription factors suggest that the third model may most accurately describe mammalian origins. However, no convincing evidence has been presented thus far directly linking the specific binding of a known transcription factor to the components of the cellular DNA synthetic apparatus and the subsequent enhanced utilization of a mammalian replication origin by this apparatus. In addition, it has been recognized that the ability of specific transcription factors to stimulate transcription does not correlate with their ability to enhance the initiation of DNA synthesis. The second and third models described in the preceding paragraph, predict that transcription factor binding to an origin of replication facilitates the initiation

of DNA synthesis by enhancing unwinding of the DNA duplex at the DUE. This implies that the transcription factor binding site must be either adjacent to or relatively near the DUE. Inspection of the transcription factor binding domains associated with the SV40 origin of replication reveal their presence near the early palindrome. The fourth model suggests that transcription factor binding to an element within a replication origin acts directly to activate a replication initiator protein; implying that the factor's binding site must be adjacent to the core origin sequence recognized by the initiator proteins.

V. WHAT WE DON'T KNOW YET ABOUT DNA SYNTHESIS

Although a wealth of information is available on the proteins, genes and molecular mechanisms involved in the replication of DNA there still are many unanswered questions regarding this pivotal process in the life cycle of a cell. Also, many of the most difficult questions deal with the replication process mediated in eukaryotic cells. Our intention is not to list every unanswered question, but to describe some of the most important problems remaining to be solved; while providing the reader some feeling for the complexity of the issues.

One of the most fundamental questions is whether the replication proteins move along the DNA strand or whether the DNA helix actually moves through a stationary replication apparatus. If the replication apparatus where to move along the double helix, the products of replication would undoubtedly move out from a rotating complex and could conceivably become entangled with one another. If however, the replication apparatus were stationery, then the two daughter strands could become segregated from one another by being directed toward two adjacent sections of the nucleus.

A second question centers on how the replication proteins are organized within the living cell. One view suggests that a functional DNA synthetic apparatus is assembled at a replication origin during S-phase, while at other points within the cell cycle the replication apparatus is disassembled into its component subunits. An alternate view suggests that these replication proteins remain assembled within the actively cycling cells, but only become activated at the onset of the S-phase.

Another question focuses upon the potential roles of the nuclear architecture in the organization of the DNA replication apparatus, the DNA strands, and regulation of the activity of the protein components making-up the replication apparatus. For example, there is a model proposed by Tubo and Berezeny (1987) that suggests that actively cycling cells contain an assembled DNA synthetic apparatus that is organized into clusters which bind around nuclear matrix filaments. Such a model favors the concept that replicating DNA moves through a stationary DNA synthetic apparatus.

A fourth question pertains to the mechanism used by cells to determine when the replication process is complete. In essence, how does the cell know that the replication of a DNA strand was completed. In the case of mammalian cells, the presence of multiple origins of replication on a single chromosomal length of DNA, typically results in one replication fork moving in one direction and ultimately colliding with a second replication fork moving in the opposite direction. But the question then becomes, why don't the origins giving rise to these replication forks initiate the next round of replication on these daughter strands prior to conclusion of the S-phase and before cell division can segregate each of the daughter DNA molecules into new daughter cells?

Precisely how and when the methylation of DNA occurs is another fundamentally important question. One model suggests that methylation of DNA occurs when a DNA methlytransferase transfers a methyl group to DNA at some finite time following the completion of the synthesis of the new DNA strand. Another model suggests that methylation occurs very shortly after creation of the nascent strand, and potentially before the strand leaves the replication apparatus.

Another fundamental observation pertaining to the replication process in mammalian cells is that the fidelity with which the replication process is carried-out results in an estimated single mutation per 10^{10} nucleotides copied. This observation has led many researchers to ask, "What makes the DNA synthetic process so faithful? One part of the answer must be that the newly synthesized DNA strands are monitored for nucleotide

mispairing. However, our understanding of how this monitoring actually occurs, and how the subsequent events leading to the correction of nucleotides mismatches is regulated remains obscure. Work performed in an attempt to address this issue invariably led to the following question.

How are DNA repair and replication linked? At some level, the proteins responsible for detecting and correcting DNA mismatches must scan and then interact directly with the newly synthesized DNA strand to correct nucleotide mispairings that occurred during the replication process. This must occur prior to the new daughter strand becoming imprinted with the parental methylation pattern. In addition, the nucleotide and base excision repair processes appear to share a number of the same proteins as the replication process. As an example, two of these proteins in mammalian cells are DNA polymerase delta and PCNA. Precisely how these proteins intermesh with the other repair proteins remains unclear, and how, as a group, they function to mediate the repair processes either while the replication process is occurring or after it is completed has still to be determined.

One of the truly significant motivations for studying the DNA synthetic process in mammalian cells stems from the question, "Does the DNA replication machinery of tumor cells differ from that of normal cells?" If it does, then the implications of having this information and utilizing it for advancing the health-care and well-being of people around the world become enormous. First, identification of specific differences in the component structure and activity of the DNA synthetic apparatus of normal and tumor cells could lead to highly specific tumor cell markers that could greatly facilitate the early detection of tumors (e.g., prior to the presence of symptoms within the individual). Second, identification of the biochemical basis for these differences in structure and activity would very likely point toward new biochemical targets that could be, in themselves, fundamental hallmarks of malignancy, and therefore new targets for highly specific therapies that could truly target only the cancer cell.

Clearly, these are only a few of the important questions associated with the replication process. Some of these are fundamental basic science questions that would provide

clearer insights into the replication process. Still other questions have direct bearing on long-term human health and improving the well being of families and friends around the world. There are essentially an enormous number of questions that are only limited by the readers imagination. The shape of those questions will help determine what we learn about the basic biochemical process of DNA replication, and how well we will use that information to help save lives and shape our future as a species.

Acknowledgments

The authors wish to thank the editors for the honor of having been selected to contribute to this very special and important project. The National Institutes of Health/National Cancer Institute and the Department of Defense Breast Cancer Research Program are also recognized for their continued support of the authors DNA replication related research projects.